

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:11:13 ; Search time 126 Seconds
(without alignments)

1004.613 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278

Sequence: 1 MKKITGILLLLAVILSAC.....EMSYTAQLVRLTEYFAKIAK 448

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

1: geneseqp1980s.*

2: geneseqp1990s.*

3: geneseqp2000s.*

4: geneseqp2001s.*

5: geneseqp2002s.*

6: geneseqp2003as.*

7: geneseqp2003bs.*

8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2278	100.0	448	5	AAM50664 Streptoco
2	1656.5	72.7	336	5	AAM50665 Streptoco
3	1656.5	72.7	336	5	AAM50639 Streptoco
4	1655.5	72.7	336	5	ABP29960 Streptoco
5	1655.5	72.7	336	6	ABU46455 Protein e
6	1655.5	72.7	345	5	ABP29106 Streptoco
7	1652.5	72.5	336	2	AAR56486 Streptoco
8	1652.5	72.5	336	4	AAY85681 Plasmin r
9	1564.5	68.7	336	5	ABP30758 Streptoco
10	1559.5	68.5	336	5	AAM50667 Streptoco
11	1559.5	68.5	336	5	AAM50641 Streptoco
12	1557.5	68.4	336	5	AAM50666 Streptoco
13	1557.5	68.4	336	5	AAM50640 Streptoco
14	1535	67.4	335	6	ABU02516 S. pneumo
15	1535	67.4	359	4	AAY37576 Streptoco
16	1535	67.4	359	6	ABU46262 Protein e
17	1534.5	67.4	336	5	AAM50669 Streptoco
18	1534.5	67.4	336	5	AAM50643 Streptoco
19	1534	67.3	359	4	AAY38000 Streptoco
20	1530.5	67.2	336	5	AAM50668 Streptoco
21	1530.5	67.2	336	5	AAM50642 Streptoco
22	1521	66.8	337	6	ABU44330 Protein e
23	1491	65.5	333	2	AAM55089 Streptoco
24	1491	65.5	333	5	ABP54583 S. pneumo
25	1491	65.5	333	7	ADC45135 S. pneumo

ALIGNMENTS

RESULT 1

AAM50664

ID AAM50664 standard; protein; 448 AA.

XX AC AAM50664;

XX AC

DT 29-AUG-2003 (revised)

DT 08-APR-2002 (first entry)

XX DT

XX DE Streptococcus GapC multiepitope fusion polyGap4.

XX KW PolyGap4; GapC; plasmin binding protein; epitope; infection; vaccine;

XX KW immunisation; mastitis; therapy.

XX OS Streptococcus dysgalactiae.

XX OS Streptococcus agalactiae.

XX OS Streptococcus parauberis.

XX OS Chimeric.

XX PN WO200196379-A2.

XX PD 20-DEC-2001.

XX PF 11-JUN-2001; 2001WO-CA000836.

XX PR 12-JUN-2000; 2000US-0211247P.

XX PA (UYSA-) UNIV SASKATCHEWAN.

XX PI Potter AA, Perez-Casal J, Fontaine M;

XX DR WPI; 2002-098051/13.

XX DR N-PSDB; ABA91327.

XX PT Novel GapC multiple epitope fusion polypeptide comprising antigenic

XX PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.

XX PT parauberis, or S. iniae GapC protein useful for treating mastitis in

XX PT vertebrates.

XX PS Claim 9; Fig 6A-C; 116pp; English.

XX CC The present sequence is that of a novel multiple epitope fusion protein,

XX CC designated PolyGap4, comprising the entire amino acid sequence of the

XX CC Streptococcus dysgalactiae GapC plasmin binding protein in addition to

XX CC unique amino acid sequences from the Streptococcus parauberis and

XX CC Streptococcus agalactiae GapC proteins. The multiple epitope protein is

XX CC produced in host cells transformed with an expression vector comprising a

XX CC chimeric gene (see ABA91327) encoding the protein. PolyGal4 is an example

Abb55601 Lactococ
Aau35254 Enterococ
Abu29370 Protein e
Abb53868 Lactococ
Adc95142 E. faeciu
Abu25342 Protein e
Abb48810 Listeria
Abu32568 Protein e
Abu23595 Protein e
Abu43810 Protein e
Abu37737 Protein e
Abp81005 N. gonorr
Abu37077 Protein e
Aag83058 S. epider
Aau42966 Protein e
Abp40668 Staphyloc
Aau36947 Staphyloc
Aau34122 Staphyloc
Abj19195 Pathogen
Abu16073 Protein e

26 1357.5 59.6 336 5 ABB55601
27 1292 56.3 333 4 AAU35254
28 1282 56.3 333 6 ABU29370
29 1275 56.0 336 5 ABB53868
30 1270 55.8 333 7 ADC95142
31 1224.5 53.8 335 6 ABU25342
32 1205.5 52.9 336 5 ABB48810
33 1205.5 52.9 336 6 ABU32568
34 1203 52.8 334 6 ABU23595
35 1166 51.2 335 6 ABU43810
36 1145 50.3 334 6 ABU37737
37 1141 50.1 357 6 ABP81005
38 1141 50.1 357 6 ABU37077
39 1138.5 50.0 336 4 AAG83058
40 1138.5 50.0 336 6 ABU42966
41 1138.5 50.0 348 5 ABP40668
42 1132.5 49.7 336 4 AAU36947
43 1132.5 49.7 336 4 AAU34122
44 1132.5 49.7 336 6 ABJ19195
45 1132.5 49.7 336 6 ABU16073

of novel GapC multiple epitope fusion proteins of the invention that comprise epitopes from 1 or more of *S. dysgalactiae*, *S. agalactiae*, *S. parauberis*, *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-69). The multiple epitope fusion proteins are used in claimed vaccines for treating or preventing a bacterial infection in a vertebrate, especially a streptococcal infection, and particularly mastitis. They are also used in claimed methods of detecting *Streptococcus* antibodies. The multiple epitope protein is capable of eliciting broad immunity against a variety of streptococcal infections while minimising the number of antigens present in the final formulation and concomitantly reducing production costs. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 448 AA;

Query Match 100.0%; Score 2278; DB 5; Length 448;
 Best Local Similarity 100.0%; Pred. No. 1.8e-180;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKITGIILLALLVILSACQANYGSMVKVINGFGRIQRLAPRRIONVEGVETRIN 60
 DB 1 MKKITGIILLALLVILSACQANYGSMVKVINGFGRIQRLAPRRIONVEGVETRIN 60

QY 61 DLTPNMLAHLIKYDTTQGRFDGTVEVKEGGFEVNGNFIKVSARDPENIDWATGVIV 120
 DB 61 DLTPNMLAHLIKYDTTQGRFDGTVEVKEGGFEVNGNFIKVSARDPENIDWATGVIV 120

QY 121 LEALGTVKDGDPVNGKTIKVSARDPENIDWATGVIVLEIDGTVEVKEGGFEV 180
 DB 121 LEALGTVKDGDPVNGKTIKVSARDPENIDWATGVIVLEIDGTVEVKEGGFEV 180

QY 181 GQFVKVSAREPANIDWATGVIVLEATSFPAKKEAEKHLHANGAKKVVITAPGNDV 240
 DB 181 GQFVKVSAREPANIDWATGVIVLEATSFPAKKEAEKHLHANGAKKVVITAPGNDV 240

QY 241 KTVFVTHDILDTGETVVISGASCTNCLAPMAKALHDAFGIQKGLMTTHAYTGDQML 300
 DB 241 KTVFVTHDILDTGETVVISGASCTNCLAPMAKALHDAFGIQKGLMTTHAYTGDQML 300

QY 301 DPHRGGLRRARAGAAIVNSTGAARAKAIGLIVPELNGKLDGAAQRPVPTGVTGLV 360
 DB 301 DPHRGGLRRARAGAAIVNSTGAARAKAIGLIVPELNGKLDGAAQRPVPTGVTGLV 360

QY 361 TLDKNVSDVINAAKKAASNDSFGYTEDPIVSSDIVGVSYGSLFDATQTKWMEVDGSLV 420
 DB 361 TLDKNVSDVINAAKKAASNDSFGYTEDPIVSSDIVGVSYGSLFDATQTKWMEVDGSLV 420

QY 421 KVVSWYDNEMSYTAQLVLTLEYFAKIAK 448
 DB 421 KVVSWYDNEMSYTAQLVLTLEYFAKIAK 448

RESULT 2
 AAM50665
 ID AAM50665 standard; protein; 336 AA.

XX AC AAM50665;

XX DT 08-APR-2002 (first entry)

DE Streptococcus dysgalactiae gapC plasmin binding protein DysGapC.

KW DysGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
 KW immunisation; mastitis; therapy.

OS Streptococcus dysgalactiae.

XX WO2001196379-A2.

XX 20-DEC-2001.

XX 11-JUN-2001; 2001WO-CA000836.

XX 12-JUN-2000; 2000US-0211247P.

XX (UYSA-) UNIV SASKATCHEWAN.
 PA Potter AA, Perez-Casal J, Fontaine M;
 PI WPI; 2002-098051/13.
 DR N-PSDB; ABA91328.
 XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
 PT determinant of *Streptococcus dysgalactiae*, *S. agalactiae*, *S. uberis*, *S.*
 PT *parauberis*, or *S. iniae* GapC protein useful for treating mastitis in
 PT vertebrates.
 XX Claim 8; Fig 1A-B; 116pp; English.
 XX The present sequence is that of the GapC plasmin binding protein,
 CC DysGapC, of *Streptococcus dysgalactiae* ATCC 43078, an isolate from a case
 CC of bovine mastitis. The invention relates to novel GapC multiple epitope
 CC fusion proteins that comprise epitopes from 1 or more of *Streptococcus*
 CC *dysgalactiae*, *Streptococcus agalactiae*, *Streptococcus parauberis*,
 CC *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-69). A claimed
 CC example is PolyGap4 (see AAM50664). Expression vectors and host cells for
 CC production of the multiple epitope fusion proteins are provided. The
 CC multiple epitope proteins are used in claimed vaccines for treating or
 CC preventing a bacterial infection in a vertebrate, especially a
 CC streptococcal infection, and particularly mastitis. They are also used in
 CC claimed methods of detecting *Streptococcus* antibodies. The multiple
 CC epitope proteins are capable of eliciting broad immunity against a
 CC variety of streptococcal infections while minimising the number of
 CC antigens present in the final formulation and concomitantly reducing
 CC production costs
 XX Sequence 336 AA;

Query Match 72.7%; Score 1656.5; DB 5; Length 336;
 Best Local Similarity 79.6%; Pred. No. 4.9e-129;
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIQRLAPRRIONVEGVETRINDTPNMLAHLIKYDTTQGRFDGTVEV 87
 DB 1 MVVKVINGFGRIQRLAPRRIONVEGVETRINDTPNMLAHLIKYDTTQGRFDGTVEV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATGVIVLEALGTVKDGDPVNGKFIKVSAR 147
 DB 61 KEGGFVNGNFIKVSARDPE-----NIDWATGVIVLE----- 81

QY 148 KDPEQIDWATGVIVLEIDGTVEVKEGGFEVNGQFVKVSAREPANIDWATGVIVLE 207
 DB 82 -----NIDWATGVIVLE 95

QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGGNDVKTVPNTNHDILDTGETVISGASCTTN 267
 DB 96 ATGFFAKKEAEKHLHANGAKKVVITAPGGNDVKTVPNTNHDILDTGETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPHERGDLRRARAGAAIVNSTGAA 327
 DB 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPHERGDLRRARAGAAIVNSTGAA 215

QY 328 KAIGLIVPELNGKLDGAAQRPVPTGVTGLVITLDKNVSDVINAAKKAASNDSFGYTE 387
 DB 216 KAIGLIVPELNGKLDGAAQRPVPTGVTGLVITLDKNVSDVINAAKKAASNDSFGYTE 275

QY 388 DPVSSDIVGVSYGSLFDATQTKWMEVDGSLVKKVSWYDNEMSYTAQLVLTLEYFAKIA 447
 DB 276 DPVSSDIVGVSYGSLFDATQTKWMEVDGSLVKKVSWYDNEMSYTAQLVLTLEYFAKIA 335

QY 448 K 448
 DB 336 K 336

RESULT 3
 AAM50639

II	AA050639	standard; protein; 336 AA.	QY	328	KAIGVIVPELNGKLDGAAQRPVPTGVS	TELVWTLDKNVSVDEINAAKAAASND	SFGYTE	387
XX	AA050639;		DB	216	KAIGVIVPELNGKLDGAAQRPVPTGVS	TELVWTLDKNVSVDEINAAKAAASND	SFGYTE	275
XX	04-APR-2002	(first entry)	QY	388	DPVSSDIVGVSYGSLFDATQTKVMEVD	GSQLVKVVSWYDNEMSYTAQLVRLT	LEYFAKTA	447
XX	Streptococcus dysgalactiae	gapC plasmid binding protein.	DB	276	DPVSSDIVGVSYGSLFDATQTKVMEVD	GSQLVKVVSWYDNEMSYTAQLVRLT	LEYFAKTA	335
XX	GapC; plasmid-binding protein; DysgalGapC; infection; mastitis; vaccine;		QY	448	K	448		
XX	diagnosis; therapy.		DB	336	K	336		
XX	Streptococcus dysgalactiae.							
XX	WO200196381-A2.							
XX	20-DEC-2001.							
XX	11-JUN-2001;	2001WO-CA000838.						
XX	12-JUN-2000;	2000US-0211022P.						
XX	(UYSA-) UNIV SASKATCHEWAN.							
XX	Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;							
XX	WPI; 2002-130725/17.							
XX	N-PSDB; ABA91248.							
XX	Novel isolated GapC protein of Streptococcus dysgalactiae, S.agalactiae,							
XX	S.uberis, S.paraberis, or S.iniae, useful as vaccine component for							
XX	treating streptococcal infection which causes mastitis in vertebrates.							
XX	Claim 1(a); Fig 1A-B; 107pp; English.							
XX	The present sequence is that of the GapC plasmid binding protein							
XX	(DysgalGapC) of Streptococcus dysgalactiae ATCC 43078, a clinical isolate							
XX	from bovine mastitis. It is encoded by the GapC gene given in ABA91248.							
XX	GapC protein, which has no signal sequence or membrane anchor domain, is							
XX	capable of eliciting an immune response in a vertebrate. The invention							
XX	provides the GapC genes and proteins of 5 Streptococcus species, as well							
XX	as recombinant vectors, host cells and vaccine compositions comprising							
XX	GapC polynucleotides or proteins. The vaccines are used to treat or							
XX	prevent a bacterial infection, especially a streptococcal infection, and							
XX	mastitis in particular (claimed). GapC proteins are also used in claimed							
XX	methods for detecting GapC antibodies, and to raise antibodies that are							
XX	used in claimed methods for detecting GapC proteins							
XX	Sequence 336 AA;							
XX	Query Match	72.7%; Score 1656.5; DB 5; Length 336;						
XX	Best Local Similarity	79.6%; Pred. No. 4.9e-129;						
XX	Matches 335; Conservative	0; Mismatches 1; Indels 85; Gaps 1;						
QY	28	MVVKVINGFGRIGRLAFRIQNVGEVETRINDLTPNMLAHLKYDITQGRFDGTV	EV	87				
DB	1	MVVKVINGFGRIGRLAFRIQNVGEVETRINDLTPNMLAHLKYDITQGRFDGTV	EV	60				
QY	88	KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEALEGTVEVKDGGFDVNGKFIKVS	AE	147				
DB	61	KEGGFEVNGNFIKVSARDPE-----		81				
QY	148	KDPEQIDWATDGVETVLEIDGTVEVKEGGFEVNGFVKVSAEREPANIDWATDGVETV	LE	207				
DB	82	-----NIDWATDGVETVLE		95				
QY	208	ATGFFAKAEAKHLHANGAKKVITAPGGNDVKTVFNTNHDITDTEVIGASCTTN		267				
DB	96	ATGFFAKAEAKHLHANGAKKVITAPGGNDVKTVFNTNHDITDTEVIGASCTTN		155				
QY	268	CLAPMAKALHDAFGIQKGLMTTHAYTGDQWILDGPHRGGDLRRARAGANIVPNSTG	AA	327				
DB	156	CLAPMAKALHDAFGIQKGLMTTHAYTGDQWILDGPHRGGDLRRARAGANIVPNSTG	AA	215				

Sequence 336 AA;

Query Match 72.7%; Score 1655.5; DB 5; Length 336;
Best Local Similarity 79.3%; Pred. No. 5.9e-129;
Matches 334; Conservative 1; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLTDPMMLAHLKYDTTQGRPDGTVEV 87
Db 1 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLTDPMMLAHLKYDTTQGRPDGTVEV 60

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGDFVNGKFIKVS 147
Db 61 KEGGFEVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEV 207
Db 82 -----NIDWATDGVIEV 95

QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTEVIGASCTTN 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTEVIGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDPHRRGGDLRRARAGAAIVPNSTGAA 327
Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDPHRRGGDLRRARAGAAIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLKXNSVDEINAAKKAASNDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLKXNSVDEINAAKKAASNDSFGYTE 275

QY 388 DPIVSSDIVGVSYGSLFDATQTKVMEVDGSQLVKVWSYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVGVSYGSLFDATQTKVMEVDGSQLVKVWSYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448
Db 336 K 336

RESULT 5
ABU46455
ID ABU46455 standard; protein; 336 AA.
XX AC ABU46455;
XX 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #31982.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Streptococcus pyogenes.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Irawick JB, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA50325.
XX New antisense nucleic acids, useful for identifying proteins or screening

for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 74379; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 336 AA;

Query Match 72.7%; Score 1655.5; DB 6; Length 336;
Best Local Similarity 79.3%; Pred. No. 5.9e-129;
Matches 334; Conservative 1; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLTDPMMLAHLKYDTTQGRPDGTVEV 87
Db 1 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLTDPMMLAHLKYDTTQGRPDGTVEV 60

QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGDFVNGKFIKVS 147
Db 61 KEGGFEVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEV 207
Db 82 -----NIDWATDGVIEV 95

QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTEVIGASCTTN 267
Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTEVIGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDPHRRGGDLRRARAGAAIVPNSTGAA 327
Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDPHRRGGDLRRARAGAAIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLKXNSVDEINAAKKAASNDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLKXNSVDEINAAKKAASNDSFGYTE 275

QY 388 DPIVSSDIVGVSYGSLFDATQTKVMEVDGSQLVKVWSYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVGVSYGSLFDATQTKVMEVDGSQLVKVWSYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448

[illegible]

CC plasminogen activator (PA) so allows a reduction in dose, and thus lowers
 CC the risk of bleeding, and may prevent reocclusion of blood vessels. The
 CC protein may be coupled to a fibrin-specific monoclonal antibody to
 CC provide targeting to clots. The plasmin receptor may be useful in human
 CC or veterinary medicine, for treatment of thrombosis and pulmonary
 CC embolism, and for solubilising clots in catheters or shunts. (Updated on
 XX 25-MAR-2003 to correct PF field.)
 XX SQ Sequence 336 AA;

Query Match 72.5%; Score 1652.5; DB 2; Length 336;
 Best Local Similarity 79.1%; Pred. No. 1e-128;
 Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;
 QY 28 MVKVGINGFGRIGRLAFRRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTV 87
 Db 1 MVKVGINGFGRIGRLAFRRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTV 60
 QY 88 KEGGFVNGNFIKVSARDPENIDWATDGV EIVLEALEGTV EKVKGDFVNGKFIKVS 147
 Db 61 KEGGFVNGNFIKVSARDPE----- 81
 QY 148 KDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207
 Db 82 -----NIDWATDGV EIVLE 95
 QY 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTGTEVIGASCTTN 267
 Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTGTEVIGASCTTN 155
 QY 268 CLAPMAKALHDAPGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 327
 Db 156 CLAPMAKALHDAPGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLTKNVSVDINAAKKAASNDSPGYTE 387
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLTKNVSVDINAAKKAASNDSPGYTE 275
 QY 388 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
 Db 276 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 8
 AAY85681
 ID AAY85681 standard; protein; 336 AA.
 XX

AC AAY85681;

DT 13-FEB-2001 (first entry)

XX Streptococcal plasmin receptor amino acid sequence.

XX Plasmin receptor; plr; immune response; vaccination; antibacterial;

XX Streptococcal infection; antihelminthic.

OS Streptococcus pyogenes.

PN US6136323-A.

XX 24-OCT-2000.

XX 11-JUL-1994; 94US-00273247.

XX 29-MAR-1989; 89US-00330849.

PR 16-MAY-1990; 90US-00524411.

XX 10-AUG-1992; 92US-00928462.

XX (UYFL) UNIV FLORIDA RES FOUND INC.

XX

PI Von Mering G, Broder C, Lottenberg R, Boyle MDP;

XX WPI; 2001-006210/01.

DR N-PSDB; AAC66140.

XX

PT Raising an immune response in a mammal, especially for vaccination
 PT against group A streptococcal infections, comprises administering an
 PT isolated streptococcal plasmin receptor protein.

XX Claim 2; Col 27-30; 17pp; English.

XX

CC Polynucleotide sequence AAC66140 encodes the Streptococcal plasmin
 CC receptor protein AAY85681. The protein is used in a method for raising an
 CC immune response in a mammal. The method comprises administering the
 CC plasmin receptor protein. The method is useful as a vaccination against
 CC group A Streptococcal infections and potentially against a broad range of
 CC infections associated with pathogens expressing glyceraldehyde-3-
 CC phosphate dehydrogenase (GAPDH) or GAPDH-like proteins at their cell
 CC surfaces. The vaccine has antibacterial and antihelminthic activity

XX SQ Sequence 336 AA;

Query Match 72.5%; Score 1652.5; DB 4; Length 336;
 Best Local Similarity 79.1%; Pred. No. 1e-128; 1; Indels 85; Gaps 1;
 Matches 333; Conservative 2; Mismatches 1;
 QY 28 MVKVGINGFGRIGRLAFRRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTV 87
 Db 1 MVKVGINGFGRIGRLAFRRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTV 60
 QY 88 KEGGFVNGNFIKVSARDPENIDWATDGV EIVLEALEGTV EKVKGDFVNGKFIKVS 147
 Db 61 KEGGFVNGNFIKVSARDPE----- 81
 QY 148 KDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207
 Db 82 -----NIDWATDGV EIVLE 95
 QY 208 ATSFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTGTEVIGASCTTN 267
 Db 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTGTEVIGASCTTN 155
 QY 268 CLAPMAKALHDAPGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 327
 Db 156 CLAPMAKALHDAPGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLTKNVSVDINAAKKAASNDSPGYTE 387
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLTKNVSVDINAAKKAASNDSPGYTE 275
 QY 388 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
 Db 276 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 9

ABP30758

ID ABP30758 standard; protein; 336 AA.

XX

AC ABP30758;

XX

DT 02-JUL-2002 (first entry)

XX

XX Streptococcus polypeptide SEQ ID NO 10692.

DE

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 XX group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus agalactiae.
 XX OS
 XX PN
 XX WO200234771-A2.
 XX PD
 XX 02-MAY-2002.
 XX
 XX 29-OCT-2001; 2001WO-GB004789.
 XX PF
 XX 27-OCT-2000; 2000GB-00026333.
 XX PR
 XX 24-NOV-2000; 2000GB-00028727.
 XX PR
 XX 07-MAR-2001; 2001GB-00005640.
 XX
 XX (CHIR-) CHIRON SPA.
 XX PA
 XX (GENO-) INST GENOMIC RES.
 XX
 XX Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
 XX PI
 XX Tettelin H;
 XX
 XX WPI; 2002-352536/38.
 XX DR
 XX N-PSDB; ABN71389.
 XX
 XX New Streptococcus protein for the treatment or prevention of infection or
 XX PT
 XX disease caused by Streptococcus bacteria, such as meningitis, and for
 XX PT
 XX detecting a compound that binds to the protein.
 XX
 XX Claim 1; Page 4179; 4525pp; English.
 XX
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 XX CC
 XX streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 XX CC
 XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 XX CC
 XX the specification. The proteins have antibacterial and antiinflammatory
 XX CC
 XX activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 XX CC
 XX antibodies that bind (I) are used in the manufacture of medicaments for
 XX CC
 XX the treatment or prevention of infection or disease caused by
 XX CC
 XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 XX CC
 XX Nucleic acids encoding (I) are used to detect Streptococcus in a
 XX CC
 XX biological sample. (I) is used to determine whether a compound binds to
 XX CC
 XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 XX CC
 XX used as a vaccine or diagnostic composition. The disease caused by
 XX CC
 XX Streptococcus that is prevented or treated may be meningitis. Nucleic
 XX CC
 XX acid encoding (I) may be used to recombinantly produce (I) and may be
 XX CC
 XX used in gene therapy. Antibodies to (I) are used for affinity
 XX CC
 XX chromatography, immunoassays, and distinguishing/identifying
 XX CC
 XX Streptococcus proteins
 XX SQ
 XX Sequence 336 AA;
 XX
 XX Query Match 68.7%; Score 1564.5; DB 5; Length 336;
 XX Best Local Similarity 74.1%; Pred. No. 2.1e-121;
 XX Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;
 XX
 XX 28 MVVKVINGFGFRIGRLAFRRIONVGEVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87
 XX Db
 XX 1 MVVKVINGFGFRIGRLAFRRIONVGEVETRINDLTDPNMLAHLKYDTTQGRF----- 54
 XX
 XX 88 KEGGFEVNGNFTKVSARDPENIDMATDGVIEVLEALGCTVEVKDGGFDVNGKFIKVSAB 147
 XX Qy
 XX 55 ----- 54
 XX
 XX 148 KDPEQIDMATDGVIEVLEIDGTVEKGGFEVNGOPVKVSAREPANIDMATDGVIEVLE 207
 XX Qy
 XX 55 -----DGTVEKGGFEVNGQFVKVSAREPANIDMATDGVIEVLE 95
 XX Db
 XX
 XX 208 ATSFPAKKEAAEKHLHANGAKKVIITAPGNDVKTIVFNTNHDIDGTETVTSASCTTN 267
 XX Qy
 XX 96 ATGFFPAKKEAEOHIEHNGAKKVIITAPGNDVKTIVFNTNHDIDGTETVTSASCTTN 155
 XX Db
 XX 268 CLAPMAKALHDAPGIQKGLMTTIHAYTQDMILDGPHRGDGLRRARAGAAIVPNSTGAA 327
 XX Qy
 XX 156 CLAPMAKALQDNFQVQKGLMTTIHAYTQDMILDGPHRGDGLRRARAGAAIVPNSTGAA 215
 XX Db
 XX 328 KAIGLVIPELNGKLDGAQRVPVPTGVTSELVTLDKNVSVDENAAKAAANDSFGYTE 387
 XX Qy

Db 216 KAIGLVIPELNGKLDGAQRVPVPTGVTSELVTLDKNVSVDENAAKAAANDSFGYTE 275
 Qy 388 DPVSSDIVGVSGLSFLPDATQTKMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
 Db 276 DPVSSDIVGVSGLSFLPDATQTKVQTVGDNQLVKVSVWYDNEMSYTSQLVRLTLEYFAKIA 335
 Qy 448 K 448
 Db 336 K 336
 RESULT 10
 AAM50667
 ID AAM50667 standard; protein; 336 AA.
 XX
 XX AAM50667;
 XX AC
 XX 08-APR-2002 (first entry)
 XX DT
 XX Streptococcus uberis gapC plasmin binding protein UberGapC.
 XX DE
 XX UberGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
 XX KW immunisation; mastitis; therapy.
 XX KW
 XX Streptococcus uberis.
 XX OS
 XX WO200196379-A2.
 XX PN
 XX 20-DEC-2001.
 XX PD
 XX 11-JUN-2001; 2001WO-CA000836.
 XX PF
 XX 12-JUN-2000; 2000US-0211247P.
 XX PR
 XX (UYSA-) UNIV SASKATCHEWAN.
 XX PA
 XX Porter AA, Perez-Casal J, Fontaine M;
 XX PI
 XX WPI; 2002-098051/13.
 XX DR
 XX N-PSDB; ABA91330.
 XX DR
 XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
 XX PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
 XX PT parauberis, or S. iniae GapC protein useful for treating mastitis in
 XX PT vertebrates.
 XX
 XX Claim 8; Fig 3A-B; 116pp; English.
 XX
 XX The present sequence is that of the GapC plasmin binding protein,
 XX CC
 XX UberGapC, of Streptococcus uberis ATCC 9927. The invention relates to
 XX CC
 XX novel GapC multiple epitope fusion proteins that comprise epitopes from 1
 XX CC
 XX or more of Streptococcus dysgalactiae, Streptococcus agalactiae,
 XX CC
 XX Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae
 XX CC
 XX (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664).
 XX CC
 XX Expression vectors and host cells for production of the multiple epitope
 XX CC
 XX fusion proteins are provided. The multiple epitope proteins are used in
 XX CC
 XX claimed vaccines for treating or preventing a bacterial infection in a
 XX CC
 XX vertebrate, especially a streptococcal infection, and particularly
 XX CC
 XX mastitis. They are also used in claimed methods of detecting
 XX CC
 XX Streptococcus antibodies. The multiple epitope proteins are capable of
 XX CC
 XX eliciting broad immunity against a variety of streptococcal infections
 XX CC
 XX while minimising the number of antigens present in the final formulation
 XX CC
 XX and concomitantly reducing production costs
 XX
 XX Sequence 336 AA;
 XX
 XX Query Match 68.5%; Score 1559.5; DB 5; Length 336;
 XX Best Local Similarity 73.9%; Pred. No. 5.5e-121;
 XX Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
 XX
 XX 28 MVVKVINGFGFRIGRLAFRRIONVGEVETRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87
 XX Qy
 XX
 XX

Db 1 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTDNMLAHLKYDTTQGRFDGTV 60
 QY 88 KEGGFVNGNFIVKSAARDPENIDWATDGVLEALEGTVKGGFDVNGKFIKVS 147
 Db 61 KGGFVNGNFIVKSAEKDP 81
 QY 148 KDPQIDWATDGVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVLE 207
 Db 82 -----NIDWATDGVLE 95
 QY 208 ATSFPAKAAEAEKHLHANGAKKVVITAPGNDVKTVPVNTNHDILDTETVISGASCTTN 267
 Db 96 ATGFFAKAAEAEKHLHANGAKKVVITAPGDDVKTVPVNTNHDILDTETVISGASCTTN 155
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGANIVPNSGAA 327
 Db 156 CLAPMAKALQDNFGVKGLMTTHAYTGDMILDGPHRGDLRRARAGANIVPNSGAA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVVLTKNVSVDEINAAKKAANDSFY 387
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVAVLEKETSVEEINAAKKAANDSFY 275
 QY 388 DPVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
 Db 276 DPVSSDIIGMAYGSLFDATQTKVQTVGNGQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 11

AAM50641
 ID AAM50641 standard; protein; 336 AA.

AC AAM50641;

XX AAM50641;

DT 04-APR-2002 (first entry)

XX Streptococcus uberis gapC plasmin binding protein.

XX GapC; plasmin-binding protein; UberGapC; infection; mastitis; vaccine;
 diagnosis; therapy.

OS Streptococcus uberis.

XX WO200196381-A2.

XX 20-DEC-2001.

XX 11-JUN-2001; 2001WO-CA000838.

XX 12-JUN-2000; 2000US-0211022P.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;

XX WPI; 2002-130725/17.

XX N-PSDB; ABA91250.

XX Novel isolated GapC protein of Streptococcus dysgalactiae, S.agalactiae,
 PT S.uberis, S.parauberis, or S.iniae, useful as vaccine component for
 PT treating streptococcal infection which causes mastitis in vertebrates.

XX Claim 1(c); Fig 3A-B; 107pp; English.

XX The present sequence is that of the GapC plasmin binding protein
 CC (UberGapC) of Streptococcus uberis ATCC 9927, encoded by the GapC gene
 CC given in ABA91250. GapC, which has no signal sequence or membrane anchor
 CC domain, is capable of eliciting an immune response in a vertebrate. The
 CC invention provides the GapC genes and proteins of 5 Streptococcus
 CC species, as well as recombinant vectors, host cells and vaccine

CC compositions comprising GapC polynucleotides or proteins. The vaccines
 CC are used to treat or prevent a bacterial infection, especially a
 CC streptococcal infection, and mastitis in particular (claimed). GapC
 CC proteins are also used in claimed methods for detecting GapC antibodies,
 CC and to raise antibodies that are used in claimed methods for detecting
 CC GapC proteins. S. uberis is a common pathogen associated with mastitis in
 CC cattle, horse, sheep and goat
 XX
 SQ Sequence 336 AA;

Query Match 68.5%; Score 1559.5; DB 5; Length 336;

Best Local Similarity 73.9%; Pred. No. 5.5e-121;

Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTDNMLAHLKYDTTQGRFDGTV 87
 Db 1 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTDNMLAHLKYDTTQGRFDGTV 60
 QY 88 KEGGFVNGNFIVKSAERDPENIDWATDGVLEALEGTVKGGFDVNGKFIKVS 147
 Db 61 KGGFVNGNFIVKSAEKDP 81

QY 148 KDPQIDWATDGVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVLE 207
 Db 82 -----NIDWATDGVLE 95

QY 208 ATSFPAKAAEAEKHLHANGAKKVVITAPGNDVKTVPVNTNHDILDTETVISGASCTTN 267
 Db 96 ATGFFAKAAEAEKHLHANGAKKVVITAPGDDVKTVPVNTNHDILDTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGANIVPNSGAA 327
 Db 156 CLAPMAKALQDNFGVKGLMTTHAYTGDMILDGPHRGDLRRARAGANIVPNSGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVVLTKNVSVDEINAAKKAANDSFY 387
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVAVLEKETSVEEINAAKKAANDSFY 275

QY 388 DPVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 447
 Db 276 DPVSSDIIGMAYGSLFDATQTKVQTVGNGQLVKVSVWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 12

AAM50666

ID AAM50666 standard; protein; 336 AA.

XX AAM50666;

AC AAM50666;

XX 08-APR-2002 (first entry)

XX Streptococcus agalactiae gapC plasmin binding protein AgalGapC.

XX AgalGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
 immunisation; mastitis; therapy.

OS Streptococcus agalactiae.

XX WO200196379-A2.

XX 20-DEC-2001.

XX 11-JUN-2001; 2001WO-CA000836.

XX 12-JUN-2000; 2000US-0211247P.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Potter AA, Perez-Casal J, Fontaine M;


```

XX WPI; 2002-098051/13.
DR N-PSDB; ABA91329.
XX
XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
PT parauberis, or S. iniae GapC protein useful for treating mastitis in
PT vertebrates.
XX
XX Claim 8; Fig 2A-B; 116pp; English.
XX
XX The present sequence is that of the GapC plasmin binding protein,
CC AgalGapC, of Streptococcus agalactiae ATCC 27541. The invention relates
CC to novel GapC multiple epitope fusion proteins that comprise epitopes
CC from 1 or more of Streptococcus dysgalactiae, Streptococcus agalactiae,
CC Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae
CC (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664).
CC Expression vectors and host cells for production of the multiple epitope
CC fusion proteins are provided. The multiple epitope proteins are used in
CC claimed vaccines for treating or preventing a bacterial infection in a
CC vertebrate, especially a streptococcal infection, and particularly
CC mastitis. They are also used in claimed methods of detecting
CC Streptococcus antibodies. The multiple epitope proteins are capable of
CC eliciting broad immunity against a variety of streptococcal infections
CC while minimising the number of antigens present in the final formulation
CC and concomitantly reducing production costs
XX
XX SQ
XX Sequence 336 AA;
XX
XX Query Match 68.4%; Score 1557.5; DB 5; Length 336;
XX Best Local Similarity 73.9%; Pred. No. 8.1e-121;
XX Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
XX
QY 28 MVVKVINGFGRIGRLAFRRIQNVGEVETRIINDLTDENMLAHLKLYDTTQGRPDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIQNVGEVETRIINDLTDENMLAHLKLYDTTQGRF----- 54
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVSAR 147
Db 55 ----- 54
QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 55 -----DGTVKEGGEVNGQFVKVSAEREPANIDWATDGEIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVWFTNHDILDTGTETVISGASCTTN 267
Db 96 ATGFPASKEKAGQIHENGAKKVVITAPGGNDVKTVWFTNHDILDTGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPFRGGDLRRARAGAANIVPNSGTAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPFRGGDLRRARAGAANIVPNSGTAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLDKXVSVDEINAAKAAANDSFGYTE 387
Db 448 K 448
Db 336 K 336

```

RESULT 13

AAM50640

ID AAM50640 standard; protein; 336 AA.

XX

AC AAM50640;

XX

DT 04-APR-2002 (first entry)

```

XX Streptococcus agalactiae gapC plasmin binding protein.
DE
XX GapC; plasmin-binding protein; AgalGapC; infection; mastitis; vaccine;
XX diagnosis; therapy.
XX
XX Streptococcus agalactiae.
OS
XX WO200196381-A2.
PN
XX 20-DEC-2001.
PD
XX 11-JUN-2001; 2001WO-CAC000838.
PF
XX 12-JUN-2000; 2000US-0211022P.
XX
XX (UYSA-) UNIV SASKATCHEWAN.
XX
XX Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
XX
XX WPI; 2002-130725/17.
XX N-PSDB; ABA91249.
XX
XX Novel isolated GapC protein of Streptococcus dysgalactiae, S. agalactiae,
XX S. uberis, S. parauberis, or S. iniae, useful as vaccine component for
XX treating streptococcal infection which causes mastitis in vertebrates.
PT
XX Claim 1(b); Fig 2A-B; 107pp; English.
XX
XX The present sequence is that of the GapC plasmin binding protein
XX (AgalGapC) of Streptococcus agalactiae ATCC 27541, encoded by the GapC
XX gene given in ABA91249. GapC, which has no signal sequence or membrane
XX anchor domain, is capable of eliciting an immune response in a
XX vertebrate. The invention provides the GapC genes and proteins of 5
XX Streptococcus species, as well as recombinant vectors, host cells and
XX vaccine compositions comprising GapC polynucleotides or proteins. The
XX vaccines are used to treat or prevent a bacterial infection, especially a
XX streptococcal infection, and mastitis in particular (claimed). GapC
XX proteins are also used in claimed methods for detecting GapC antibodies,
XX and to raise antibodies that are used in claimed methods for detecting
XX GapC proteins. S. agalactiae is a common pathogen associated with
XX mastitis in cattle, horse, sheep and goat. It also causes septicemia,
XX meningitis, bacteraemia, impetigo, arthritis, urinary tract infections,
XX abscesses, spontaneous abortion, etc
XX
XX SQ
XX Sequence 336 AA;
XX
XX Query Match 68.4%; Score 1557.5; DB 5; Length 336;
XX Best Local Similarity 73.9%; Pred. No. 8.1e-121;
XX Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
XX
QY 28 MVVKVINGFGRIGRLAFRRIQNVGEVETRIINDLTDENMLAHLKLYDTTQGRPDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIQNVGEVETRIINDLTDENMLAHLKLYDTTQGRF----- 54
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVSAR 147
Db 55 ----- 54
QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 55 -----DGTVKEGGEVNGQFVKVSAEREPANIDWATDGEIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVWFTNHDILDTGTETVISGASCTTN 267
Db 96 ATGFPASKEKAGQIHENGAKKVVITAPGGNDVKTVWFTNHDILDTGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPFRGGDLRRARAGAANIVPNSGTAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPFRGGDLRRARAGAANIVPNSGTAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLDKXVSVDEINAAKAAANDSFGYTE 387
Db 448 K 448
Db 336 K 336

```

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVATLEKDVTVVEVNAAMKAAANDSYGYTE 275
 QY 388 DPVSSDIVGSGSLFDATQTKMEVDGSQLVKVWSYDNEMSYTAQLVRLTYEYAKIA 447
 Db 276 DPVSSDIVGSGSLFDATQTKVQTVGQNLVKVWSYDNEMSYTSQLVRLTYEYAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 14
 ABU02516
 ID ABU02516 standard; protein; 335 AA.
 XX
 AC ABU02516;
 DT 23-OCT-2003 (revised)
 DT 11-FEB-2003 (first entry)
 XX
 DE S. pneumoniae type 4 strain protein from coding region #2094.
 XX
 DE Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
 KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KW gene therapy; vaccine.
 XX
 OS Streptococcus pneumoniae; type 4 strain.
 XX
 XX WO200277021-A2.
 PN
 XX 03-OCT-2002.
 PD
 XX 27-MAR-2002; 2002WO-IB002163.
 PF
 XX 27-MAR-2001; 2001GB-00007658.
 PR
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 XX Masignani V, Tettelin H, Fraser C;
 WPI: 2003-040579/03.
 DR N-PSDB; ABX07806.
 XX
 PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
 PT ear infection.

PS Claim 1; SEQ ID NO 4188; 56pp; English.

XX The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC ABS56454. Also included are an antibody which binds one of the proteins,
 CC treating a patient by administering the protein, DNA or antibody (in a
 CC composition), a kit comprising first and second primers, which are the
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a
 CC sequence not defined in the specification, for amplifying a target
 CC sequence contained within a Streptococcus nucleic acid sequence, where
 CC the first primer is substantially complementary to the target sequence,
 CC and the second primer is substantially complementary to the complement of
 CC the target sequence, and where the parts of the primers having
 CC substantial complementarity define the termini of the target sequence to
 CC be amplified, assay comprising contacting a test compound with the
 CC protein, and determining whether the test compound binds to the protein
 CC and a Streptococcus pneumoniae bacterium, where one or more genes
 CC encoding the proteins has been rendered inactive. The proteins, nucleic
 CC acid molecules, antibody and compositions are useful as medicaments for
 CC treating or preventing a disease or infection due to streptococcus
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis

CC media or ear infection. They are also useful in developing vaccines,
 CC diagnostics and antibiotics. The methods are useful for identifying
 CC immunodominant proteins. The present sequence is one of the 2469 proteins
 CC expressed by the identified coding regions from the genomic sequence.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pot_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 335 AA;

Query Match 67.4%; Score 1535; DB 6; Length 335;
 Best Local Similarity 73.6%; Pred. No. 6e-119;
 Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
 QY 28 MVMKVGINGFGRIGRLAFRRIQNVGEVETRIINDLTDPNMLAHLKYDTTQGRFDGTV 87
 Db 1 MVMKVGINGFGRIGRLAFRRIQNVGEVETRIINDLTDPNMLAHLKYDTTQGRFDGTV 60
 QY 88 KEGGFEVNGNFVKVSAERDPENIDWATDGEIVLEALEGTVEVKDGGFDVNGKFIKVS 147
 Db 61 KEGGFEVNGKFIKVSARDPE----- 81
 QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
 Db 82 -----QIDWATDGEIVLE 95
 QY 208 ATSFPAKAEAKHHLHANGAKKVVITAPGNDVKTVVFNTHDILDTGTTVIGASCTTN 267
 Db 96 ATGFFAKAEAEKHL-KGGAKKVVITAPGNDVKTVVFNTHDVLDTGTTVIGASCTTN 154
 QY 268 CLAPMAKALHDAFGTQKGLMTTIHAYTGDMILDGHPRGDLRRARAGANIVPNSTGAA 327
 Db 155 CLAPMAKALQDNFVGVVEGLMTTIHAYTGDMILDGHPRGDLRRARAGANIVPNSTGAA 214
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVTVLIDKNVSDVINAAKAAANDSYGYTE 387
 Db 215 KAIGLVIPELNGKLDGSAQRPVPTGVTSELVAVLEKVVDEVNAAMKAAANDSYGYTE 274
 QY 388 DPVSSDIVGSGSLFDATQTKMEVDGSQLVKVWSYDNEMSYTAQLVRLTYEYAKIA 447
 Db 275 DPVSSDIVGSGSLFDATQTKVLDVKGQLVKVWSYDNEMSYTAQLVRLTYEYAKIA 334
 QY 448 K 448
 Db 335 K 335

RESULT 15

AAU37576
 ID AAU37576 standard; protein; 359 AA.
 XX
 AC AAU37576;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Streptococcus pneumoniae cellular proliferation protein #5.
 DE
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200170955-A2.
 PD 27-SEP-2001.
 XX
 XX 21-MAR-2001; 2001WO-US009180.
 PF
 XX 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.

Search completed: September 15, 2004, 10:24:29
Job time : 131 secs

```
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX PI Haselbeck R, Ohlson KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS55435.
XX
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX FS Example 3; SEQ ID NO 13169; 511pp; English.
XX
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes
XX CC themselves and the encoded proteins. The prokaryotes used are Escherichia
XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence represents an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 359 AA;
Query Match 67.4%; Score 1535; DB 4; Length 359;
Best Local Similarity 73.6%; Pred. No. 6.6e-119;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
QY 28 MVVKVINGFGRIGRLAFRRICQNVGVEVTRINDLDPNLAHLKDYDTTQGRPDGTVEV 87
Db 25 MVVKVINGFGRIGRLAFRRICQNVGVEVTRINDLDPNLAHLKDYDTTQGRPDGTVEV 84
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
Db 85 KEGGFVNGNFIKVSARDPE----- 105
QY 148 KDEQIDWATDGVLEIDGTVEVKEGGFVNGQFVKVSAREPANIDWATDGVLEIVLE 207
Db 106 -----QIDWATDGVLEIVLE 119
QY 208 ATSPFAKKEAAEKHLHANGAKKVITAPGNDVKTWVNTNHDILDTGTETVISGASCTTN 267
Db 120 ATGFPFAKKEAAEKHL-KEGAKKVITAPGNDVKIWFNTNHDVLDGTETVISGASCTTN 178
QY 268 CIAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGFHRRGGDLRRARAGAANIVPNSGTAA 327
Db 179 CIAPMAKALQDNFVGVEGLMTTTHAYTGDQMLDGFHRRGGDLRRARAGAANIVPNSGTAA 238
QY 328 KAIGLVIPELNGKLDGAQRVPVPTGVTTELVTLDKNVSVUDEINAAKKAASNDSPGYTE 387
Db 239 KAIGLVIPELNGKLDGSAQRVPTPTGVTTELVAVLEKXVTVDVNAAMKAASNESYGYTE 298
QY 388 DPVSSDDIVGVSYSIFDQTKWMEVDGSQLVKVVSVDNEMSYTAQLVRTLEYFAKIA 447
Db 299 DPVSSDDIVGVSYSIFDQTKVLDVKGQLVKVVSVDNEMSYTAQLVRTLEYFAKIA 358
QY 448 K 448
Db 359 K 359
```

this Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:00:07 ; Search time 34 Seconds
(without alignments)
680.248 Million cell updates/sec

Title: US-10-650-369-22
Perfect score: 2278
Sequence: 1 MKKITGIILLAVIILSAC.....EMSYAQLVRLTFYFAKIAK 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/aaa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/aaa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/aaa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/aaa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/aaa/PTUS_COMB.pep:*
6: /cgn2_6/prodata/2/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2278	100.0	448	US-09-878-766A-22	Sequence 22, Appl
2	1656.5	72.7	336	US-09-878-766A-12	Sequence 12, Appl
3	1652.5	72.5	336	US-07-928-462-2	Sequence 2, Appl
4	1652.5	72.5	336	US-08-273-247-2	Sequence 2, Appl
5	1559.5	68.5	336	US-09-878-766A-16	Sequence 16, Appl
6	1557.5	68.4	336	US-09-878-766A-14	Sequence 14, Appl
7	1534.5	67.4	336	US-09-878-766A-20	Sequence 20, Appl
8	1530.5	67.2	336	US-09-878-766A-18	Sequence 18, Appl
9	1491	65.5	333	US-08-961-083-54	Sequence 54, Appl
10	1491	65.5	333	US-09-536-784-54	Sequence 54, Appl
11	1282	56.3	346	US-09-134-000C-4400	Sequence 4400, Ap
12	1270	55.8	333	US-09-107-532A-4769	Sequence 4769, Ap
13	1138.5	50.0	348	US-09-134-000C-5513	Sequence 5513, Ap
14	923.5	40.5	357	US-09-134-000C-4229	Sequence 4229, Ap
15	918.5	40.3	340	US-09-634-238-233	Sequence 233, Ap
16	905.5	39.7	349	US-09-489-039A-13939	Sequence 13939, A
17	849	37.3	335	US-09-690-9	Patent No. 5290690
18	838	36.8	334	5290690-10	Patent No. 5290690
19	803	35.3	340	US-09-095-855-205	Sequence 205, App
20	803	35.3	340	US-09-205-426-205	Sequence 205, App
21	780	34.2	334	US-09-690-11	Patent No. 5290690
22	776.5	34.1	335	US-09-489-039A-7679	Sequence 7679, Ap
23	758.5	33.3	282	US-09-107-532A-4057	Sequence 4057, Ap
24	747.5	32.8	340	US-09-543-681A-8083	Sequence 8083, Ap
25	719	31.6	352	US-09-489-039A-9410	Sequence 9410, Ap
26	690	30.3	336	US-09-198-452A-662	Sequence 662, App
27	683.5	30.0	338	US-09-091-725-10	Sequence 10, Appl

28	678	29.8	268	2	US-08-997-080-186	Sequence 186, App
29	678	29.8	268	2	US-08-997-362-186	Sequence 186, App
30	678	29.8	268	3	US-09-095-855-186	Sequence 186, App
31	678	29.8	288	4	US-09-324-542-186	Sequence 186, App
32	678	29.8	288	4	US-09-205-426-186	Sequence 186, App
33	675	29.6	335	2	US-08-903-800A-6	Sequence 6, Appli
34	668.5	29.3	340	4	US-09-252-991A-21612	Sequence 21612, A
35	660	29.0	334	4	US-09-674-826B-6	Sequence 6, Appli
36	646	28.4	341	4	US-09-134-001C-3783	Sequence 3783, Ap
37	628	27.6	333	3	US-09-532-803-11	Sequence 11, Appl
38	617	27.1	333	3	US-09-532-803-9	Sequence 13, Appl
39	601.5	26.4	277	4	US-09-548-938A-13	Sequence 3378, Ap
40	542.5	23.8	442	4	US-09-540-236-3378	Sequence 11854, A
41	540.5	23.7	350	4	US-09-489-039A-11854	Sequence 6704, Ap
42	535	23.5	340	4	US-09-328-352-6704	Sequence 3470, Ap
43	532.5	23.4	417	4	US-09-540-236-3470	Sequence 6082, Ap
44	512.5	22.5	359	4	US-09-543-681A-6082	Sequence 4616, Ap
45	504.5	22.1	490	4	US-09-328-352-4616	

ALIGNMENTS

RESULT 1
US-09-878-766A-22
; Sequence 22, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
; OTHER INFORMATION: epitope fusion protein
US-09-878-766A-22

Query Match	100.0%;	Score	2278;	DB	4;	Length	448;
Best Local Similarity	100.0%;	Pred. No.	1.7e-206;				
Matches	448;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MKKITGIILLAVIILSACQANYGSGMVVKVINGFGRIGRLAFRRIQNVGVGVETRIIN	60				
Db	1	MKKITGIILLAVIILSACQANYGSGMVVKVINGFGRIGRLAFRRIQNVGVGVETRIIN	60				
Qy	61	DLTDENLAHLKYDTTQGRFDGTVEVKEGFGFVNGNFVKVSAERDENIDWATDGVIEIV	120				
Db	61	DLTDENLAHLKYDTTQGRFDGTVEVKEGFGFVNGNFVKVSAERDENIDWATDGVIEIV	120				
Qy	121	LEALEGTVEVKDGFVNGKPKIKVSAEKDPEQIDWATDGVIEIVLEIDGTVEVKEGFEVN	180				
Db	121	LEALEGTVEVKDGFVNGKPKIKVSAEKDPEQIDWATDGVIEIVLEIDGTVEVKEGFEVN	180				
Qy	181	GQFVKVSAEREPANIDWATDGVIEIVLEATSFPAKKEAAEKHLHANGAKKVVITAPGNDV	240				
Db	181	GQFVKVSAEREPANIDWATDGVIEIVLEATSFPAKKEAAEKHLHANGAKKVVITAPGNDV	240				
Qy	241	KTVVFNTHDILDTGTEVIVSGASCTTNCLAPMAKALHDAFGIQKGLMTTHAYTGDQML	300				
Db	241	KTVVFNTHDILDTGTEVIVSGASCTTNCLAPMAKALHDAFGIQKGLMTTHAYTGDQML	300				
Qy	301	DGPHRGDGLRPARAGAAINVENSTGAAKIGLVTPELNGKLDGAAQRPVPTGSGTELVV	360				

Db 301 DGPGRGDLRRARAGANIVPNSGTGAAGAIGLVIPELNGKLDGAAQRPVPTGVTVELVV 360

QY 361 TLDKNVSVDEINAAKASNDGFGYTEDPIVSSDIVGVSGLFDATQTKWMEVDGSQLV 420

Db 361 TLDKNVSVDEINAAKASNDGFGYTEDPIVSSDIVGVSGLFDATQTKWMEVDGSQLV 420

QY 421 KVVSVYDNEMSYTAQLVRLTFYFAKIAK 448

Db 421 KVVSVYDNEMSYTAQLVRLTFYFAKIAK 448

RESULT 2

US-09-878-766A-12

; Sequence 12, Application US/09878766A

; Patent No. 6660270

; GENERAL INFORMATION:

; APPLICANT: Potter, Andrew A.

; APPLICANT: Perez-Casal, Jose

; APPLICANT: Fontaine, Michael

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN

; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0057

; CURRENT APPLICATION NUMBER: US/09/878,766A

; CURRENT FILING DATE: 2001-09-10

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 12

; LENGTH: 336

; TYPE: PRT

; ORGANISM: Streptococcus dysgalactiae

US-09-878-766A-12

Query Match 72.7%; Score 1656.5; DB 4; Length 336;

Best Local Similarity 79.6%; Pred. No. 5e-148;

Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRILAFRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTV 87

Db 1 MVVKVINGFGRIGRILAFRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVVEKGGFDVNGKFIKVS 147

Db 61 KEGGFVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFVNGQFVKVSAREPANIDWATDGVLE 207

Db 82 -----NIDWATDGVLE 95

QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 267

Db 96 ATGFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGANIVPNSGTAA 327

Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGANIVPNSGTAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVTLDKNVSVDENAAKASNDGFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVTLDKNVSVDENAAKASNDGFGYTE 275

QY 388 DPVSSDIVGVSGLFDATQTKWMEVDGSQLVKNVSVYDNEMSYTAQLVRLTFYFAKIA 447

Db 276 DPVSSDIVGVSGLFDATQTKWMEVDGSQLVKNVSVYDNEMSYTAQLVRLTFYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 3

US-07-928-462-2

; Sequence 2, Application US/07928462

; Patent No. 5328996

; GENERAL INFORMATION:

; APPLICANT: Boyle, Michael D.P.

; APPLICANT: Lottenberg, Richard

; APPLICANT: Broder, Christopher C.

; APPLICANT: von Mering, Gregory O.

; TITLE OF INVENTION: Bacterial Plasmin Receptors as

; TITLE OF INVENTION: Fibrinolytic Agents

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: David R. Saliwanchik

; STREET: 2421 N.W. 41st Street, Suite A-1

; CITY: Gainesville

; STATE: FL

; COUNTRY: USA

; ZIP: 32606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/928,462

; FILING DATE: 19920810

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/524,411

; FILING DATE: 16-MAY-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/330,849

; FILING DATE: 29-MAR-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: Saliwanchik, David R.

; REGISTRATION NUMBER: 31,794

; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 904-375-8100

; TELEFAX: 904-372-5800

; INFORMATION FOR SEQ ID NO: 2:

; LENGTH: 336 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-928-462-2

Query Match 72.5%; Score 1652.5; DB 1; Length 336;

Best Local Similarity 79.1%; Pred. No. 1.2e-147;

Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRILAFRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTV 87

Db 1 MVVKVINGFGRIGRILAFRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVVEKGGFDVNGKFIKVS 147

Db 61 KEGGFVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFVNGQFVKVSAREPANIDWATDGVLE 207

Db 82 -----NIDWATDGVLE 95

QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 267

Db 96 ATGFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGANIVPNSGTAA 327

Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGANIVPNSGTAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVTLDKNVSVDENAAKASNDGFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVTLDKNVSVDENAAKASNDGFGYTE 275

QY	388	DPVSDIVGVSYGSLFDATQFKWEVDGSQLVKVSWYDNEMSTAOILVRLTLEYPAKIA	447
1	276	DPVSGDGVGVSYGSLFDATQFKWEVDGSQLVKVSWYDNEMSTAOILVRLTLEYPAKIA	335

QY	448	K	448
Db	336	K	336

```

RESULT 4
US-08-273-247-2
; Sequence 2, Application US/08273247
; Patent No. 6136323
; GENERAL INFORMATION:
; APPLICANT: Boyle, Michael D.P.
; APPLICANT: Lottenberg, Richard
; APPLICANT: Broder, Christopher C.
; APPLICANT: von Mering, Gregory O.
; TITLE OF INVENTION: Bacterial Plasmin Receptors as
; TITLE OF INVENTION: Fibrinolytic Agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,247

```

```

1  FILING DATE:
2  CLASSIFICATION: 424
3  PRIORITY APPLICATION DATA:
4  APPLICATION NUMBER: US 07/524,411
5  FILING DATE: 16-MAY-1990
6  PRIORITY APPLICATION DATA:
7  APPLICATION NUMBER: US 07/330,849
8  FILING DATE: 29-MAR-1989
9  ATTORNEY/AGENT INFORMATION:
10 NAME: Sallwanchik, David R.
11 REGISTRATION NUMBER: 31,794
12 REFERENCE/DOCKET NUMBER: UF/S&S-1
13 TELECOMMUNICATION INFORMATION:
14 TELEPHONE: 904-375-8100
15 TELEFAX: 904-372-5800
16 INFORMATION FOR SEQ ID NO: 2:
17 SEQUENCE CHARACTERISTICS:
18 LENGTH: 336 amino acids
19 TYPE: amino acid
20 TOPOLOGY: linear
21 MOLECULE TYPE: protein
22 US-08-273-247-2

```

	Query Match	72.5%;	Score 1652.5;	DB 3;	Length 336;
	Best Local Similarity	79.1%;	Prod. No. 1.2e-147;		
	Matches 333;	Conservative	2;	Mismatches 1;	Indels 85; Gaps 1;
QY	28	MVVKVINGFGRIGLAFRRIONVGEVTRINDLTDPNMLAHLKLYDTTQGRDGVTEV	87		
Db	1	MVVKVINGFGRIGLAFRRIONVGEVTRINDLTDPNMLAHLKLYDTTQGRDGVTEV	60		
QY	88	KEGGFEVNGNFIKVSARDPENIDWATDGVIEALEGTVVEKDGSGFVNGKFIKVSAAE	147		
Db	61	KEGGFEVNGNFIKVSARDPE	81		
QY	148	KDPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE	207		
Db	92	NIDWATDGVIEVLE	95		

208	ATSFPAKEAAEXHLHANGAKVVIITAPGNDVKTVFNTNHDILDGTETVISGASCTTN	267
96	ATGFPAKEAAEXHLHANGAKVVIITAPGNDVKTVFNTNHDILDGTETVISGASCTTN	155
268	CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPHRGGDLRRPAGAAANVPNSTGAA	327
156	CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPHRGGDLRRPAGAAANVPNSTGAA	215
328	KAIGLVIPELNGKLDGAAQRPVPTGSVTELVNVLTKNVSVDENMAKKAANDSFGYTE	387
216	KAIGLVIPELNGKLDGAAQRPVPTGSVTELVNVLTKNVSVDENMAKKAANDSFGYTE	275
388	DPVSSDIVGVSYGSLFDATQTKWEVDGSQLVKVYSWYDNEMSYTAQLVRLTEYFAKIA	447
276	DPVSSDIVGVSYGSLFDATQTKWEVDGSQLVKVYSWYDNEMSYTAQLVRLTEYFAKIA	335
448	K 448	
	↑	
336	K 336	
Db		

RESULT 5
US-09-878-766A-16
Sequence 16, Application US/09878766A
Patent No. 6660270
GENERAL INFORMATION:
APPLICANT: Potter, Andrew A.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
TITLE OF INVENTION: IMMUNIZATION OF MAN
AGAINST STREPTOCOCCUS UBERIS
FILE REFERENCE: 9000-0057
CURRENT APPLICATION NUMBER: US/09/878-
766A-16
CURRENT FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 16
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus uberis
US-09-878-766A-16

Query Match	68.5%	Score 1559.5	DB 4	Length 336
Best Local Similarity	73.9%	Pred. No. 7.1e-139		
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 17				
QY	28	MVVKVGVNGFGRIGRLAFRRIQNVGEGVEVTRINDLTDPNMLAHLKYDITQGREDDGTVEV	87	
DB	1	MVVKVGVNGFGRIGRLAFRRIQNVGEGVEVTRINDLTDPNMLAHLKYDITQGREDDGTVEV	60	
QY	88	KEGGEFVNGNFIKVSABERDPENIDWATDGVIEIVLEALGTVVEVKDGGFDVNGKFIKVSAAE	147	
DB	61	KDGGFEVNGNFIKVSAAEKDPE-----	81	
QY	148	KDPEQIDWATDGVIEIVLEIDGTVEKGGFEVNGQFVKVSAEREPANIDWATDGVIEIVLE	207	
DB	82	-----NIDWATDGVIEIVLE	95	
QY	208	ATSPFAKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN	267	
DB	96	ATGFFAKAAAEKHLHANGAKKVVITAPGGDDVKTVVFNTHDILDGTETVISGASCTTN	155	
QY	268	CLAPMAKALHDFAFGIQKGLMTTIHAYTGDQMILOGPHRGDGLRRARAGAAINVPNSTGAA	327	
DB	156	CLAPMAKALQNFVGKQGLMTTIHAYTGDQMILOGPHRGDGLRRARAGASINVPNSTGAA	215	
QY	328	KAIGLVIPELNGKLDGAQRVPVPTGSVTELVTWILDKNVSVDEINAAKAAASNDSFGYTE	387	
DB	216	KAIGLVIPELNGKLDGAQRVPVPTGSVTELVAVLEKETSVEEINAAKAAANDSYGYTE	275	
QY	388	DPVSSDVIQVGSLSLFDATQTKVMEVDGSQLVKVSVYDQNMESVTAQLVRTLEVFAKIA	447	

Db 276 DPVSSDIIMAYGSLFDATQTKVQVDGNQLVKVSWYDNMSYTAQLVRTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 6

US-09-878-766A-14
 ; Sequence 14, Application US/09878766A
 ; Patent No. 6660270

GENERAL INFORMATION:

; APPLICANT: Potter, Andrew A.
 ; APPLICANT: Perez-Casal, Jose
 ; APPLICANT: Fontaine, Michael
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
 ; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
 ; FILE REFERENCE: 9000-0057
 ; CURRENT APPLICATION NUMBER: US/09/878,766A
 ; CURRENT FILING DATE: 2001-09-10
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 14
 ; LENGTH: 336
 ; TYPE: PRT
 ; ORGANISM: Streptococcus agalactiae
 US-09-878-766A-14

Query Match 68.4%; Score 1557.5; DB 4; Length 336;
 Best Local Similarity 73.9%; Pred. No. 1.1e-138;
 Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRIQNVGVEVTRINDLTPNNLAHLKYDTTQGRFDGTV 87
 Db 1 MVKVGINGFGRIGRLAFRIQNVGVEVTRINDLTPNNLAHLKYDTTQGRF 54
 QY 88 KEGGFVNGNFIKVSARDPENIDWATDGV EIVLEALEGTVEVKDGGFDVNGKFIKVS 147
 Db 55 ----- 54
 QY 148 KDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207
 Db 55 -----DGTV EKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 95
 QY 208 ATSFPAKAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTVISGASCTTN 267
 Db 96 ATGFFASKEKAGQHIEHANGAKKVVITAPGNDVKTVVNTNHDILDTGTVISGASCTTN 155
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSGTAA 327
 Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSGTAA 215
 QY 328 KAIGLVIPELNGKLGAAQRPVPTGSVTELVTLDKNVSVDEINAAKKAANDSFGYTE 387
 Db 216 KAIGLVIPELNGKLGAAQRPVPTGSVTELVTLEKDVVEVNAAKKAANDSFGYTE 275
 QY 388 DPVSSDIVGVSGLFDATQTKVMEVDSQLVKVSWYDNMSYTAQLVRTLEYFAKIA 447
 Db 276 DPVSSDIVGVSGLFDATQTKVQVDGNQLVKVSWYDNMSYTAQLVRTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 7

US-09-878-766A-20
 ; Sequence 20, Application US/09878766A
 ; Patent No. 6660270

GENERAL INFORMATION:

; APPLICANT: Potter, Andrew A.
 ; APPLICANT: Perez-Casal, Jose
 ; APPLICANT: Fontaine, Michael

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
 ; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
 ; FILE REFERENCE: 9000-0057
 ; CURRENT APPLICATION NUMBER: US/09/878,766A
 ; CURRENT FILING DATE: 2001-09-10
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 20
 ; LENGTH: 336
 ; TYPE: PRT
 ; ORGANISM: Streptococcus iniae
 US-09-878-766A-20

Query Match 67.4%; Score 1534.5; DB 4; Length 336;
 Best Local Similarity 72.7%; Pred. No. 1.6e-136;
 Matches 306; Conservative 17; Mismatches 13; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRIQNVGVEVTRINDLTPNNLAHLKYDTTQGRFDGTV 87
 Db 1 MVKVGINGFGRIGRLAFRIQNVGVEVTRINDLTPNNLAHLKYDTTQGRFDGTV 60
 QY 88 KEGGFVNGNFIKVSARDPENIDWATDGV EIVLEALEGTVEVKDGGFDVNGKFIKVS 147
 Db 61 KDGGFVNGS----- 70
 QY 148 KDPEQIDWATDGV EIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGV EIVLE 207
 Db 71 -----FVKVSAEREPANIDWATDGV EIVLE 95
 QY 208 ATSFPAKAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTVISGASCTTN 267
 Db 96 ATGFFASKEKAGQHIEHANGAKKVVITAPGNDVKTVVNTNHDILDTGTVISGASCTTN 155
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSGTAA 327
 Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSGTAA 215
 QY 328 KAIGLVIPELNGKLGAAQRPVPTGSVTELVTLDKNVSVDEINAAKKAANDSFGYTE 387
 Db 216 KAIGLVIPELNGKLGAAQRPVPTGSVTELVALEKDVVEVNAAKKAANDSFGYTE 275
 QY 388 DPVSSDIVGVSGLFDATQTKVMEVDSQLVKVSWYDNMSYTAQLVRTLEYFAKIA 447
 Db 276 DAIVSSDIVGVSGLFDATQTKVQVDGNQLVKVSWYDNMSYTAQLVRTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 8

US-09-878-766A-18
 ; Sequence 18, Application US/09878766A
 ; Patent No. 6660270

GENERAL INFORMATION:

; APPLICANT: Potter, Andrew A.
 ; APPLICANT: Perez-Casal, Jose
 ; APPLICANT: Fontaine, Michael
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
 ; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
 ; FILE REFERENCE: 9000-0057
 ; CURRENT APPLICATION NUMBER: US/09/878,766A
 ; CURRENT FILING DATE: 2001-09-10
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 18
 ; LENGTH: 336
 ; TYPE: PRT
 ; ORGANISM: Streptococcus parauberis
 US-09-878-766A-18

Query Match 67.2%; Score 1530.5; DB 4; Length 336;
 Best Local Similarity 73.4%; Pred. No. 3.9e-136;


```

Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2;
Qy 28 MVVKVINGFGRIGRLAFRRIQNVGEVETRIINDLTDPNMLAHLKKYDTTQGRDFGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIQNVGEVETRIINDLTDPNMLAHLKKYDTTQGRDFGTVEV 55
Qy 88 KEGGFVNGNFKVSAERDPENIDWATDGVIEVLEALGTVGVKGGFDVNGKFKVSAE 147
Db 56 -----GTVEVKDGGFDVNGKFKVSAE 77
Qy 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 78 KDPEQ-----IDWATDGVIEVLE 95
Qy 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTWVNTNHDILDTGTETVISAASCTTN 267
Db 96 ATGFFAKKAAAEKHLHENGAKKVVITAPGGNDVKTWVNTNHDILDTGTETVISAASCTTN 155
Qy 268 CLAPMAKALHDAFGIOKGLMTTIIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 215
Qy 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTVELVTLDKNVSVDEINAAKKAASNDGSGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTVELVAVLNKETSVEEINSVMKAAANDSGYTE 275
Qy 388 DPVSSDIVGVSGLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPVSSDIVGVSGLFDATQTKVQTDVGNQLVKVSWYDNEMSYTAQLDRTLLEYFAKIA 335
Qy 448 K 448
Db 336 K 336

RESULT 9
US-08-961-083-54
; Sequence 54, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-54

Query Match
Best Local Similarity 65.5%; Score 1491; DB 3; Length 333;
Matches 302; Conservative 12; Mismatches 17; Indels 86; Gaps 2;
Qy 29 VVKVINGFGRIGRLAFRRIQNVGEVETRIINDLTDPNMLAHLKKYDTTQGRDFGTVEV 88
Db 1 VVKVINGFGRIGRLAFRRIQNVGEVETRIINDLTDPNMLAHLKKYDTTQGRDFGTVEV 60
Qy 89 EGGFEVNGNFKVSAERDPENIDWATDGVIEVLEALGTVGVKGGFDVNGKFKVSAEK 148
Db 61 EGGFEVNGKFKVSAERDPE----- 80
Qy 149 DPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLEA 208
Db 81 -----QIDWATDGVIEVLEA 95
Qy 209 TSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTWVNTNHDILDTGTETVISAASCTTN 268
Db 96 TGFFAKKEAAEKHL-KGGAKKVVITAPGGNDVKTWVNTNHDVLDGTETVISAASCTTN 154
Qy 269 LAPMAKALHDAFGIOKGLMTTIIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGA 328
Db 155 LAPMAKALQDNFVGVLMTTIIHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGA 214
Qy 329 AIGLVIPELNGKLDGAAQRPVPTGSGVTVELVTLDKNVSVDEINAAKKAASNDGSGYTE 388
Db 215 AIGLVIPELNGKLDGSAQRPVPTGSGVTVELVAVLEKNVTVDVNAAMKAASNESGYTED 274
Qy 389 PIVSSDIVGVSGLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAK 445
Db 275 PIVSSDIVGVSGLFDATQTKVLDVGKQLVKVSWYDNEMSYTAQLVRLTILRLK 331

RESULT 10
US-09-536-784-54
; Sequence 54, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 54:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-536-784-54

Query Match 65.5%; Score 1491; DB 4; Length 333;
Best Local Similarity 72.4%; Pred. No. 2.1e-132;
Matches 302; Conservative 12; Mismatches 17; Indels 86; Gaps 2;
29 VVKVGINGFGRIQRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRDGTVEVK 88
1 VVKVGINGFGRIQRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRDGTVEVK 60
89 EGGFEVNGNFIKVSARDPENIDWATDGVLEALGTVVEVKGGFDVNGKFIKVSAAK 148
61 EGGFEVNGKFIKVSARDPE----- 80
149 DPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAREPANIDWATDGVLEIVLEA 208
81 -----QIDWATDGVLEIVLEA 95
209 TSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTGTTETVISGASCTTNC 268
96 TGFFAKKEAAEKHL-KGGAKKVVITAPGGNDVKTVVFNTHDILDTGTTETVISGASCTTNC 154
269 LAPMAKALHDAPFQIGKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSGAAK 328
155 LAPMAKALQDNFGVWEGGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSGAAK 214
329 AIGLVIPELNGKLDGAAQRPVPTGTVTELVTLDKNVSVDEINAAKASNDSPGYTED 388
215 AIGLVIPELNGKLDGAAQRPVPTGTVTELVALEKRVTVDEVNAAKAASNESYGYTED 274
389 PIVSSDIVGVSGLSFDATQTKVMEVDGSQLKVVSWYDNEMSYTAQLVRLTLEYFAK 445
275 PIVSSDIVGVSGLSFDATQTKVLDVDGKLVKVVSWYDNEMSYTAQLVRLTLEYFAK 331

RESULT 11
US-09-134-000C-4400
Sequence 4400, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4400
LENGTH: 346
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-4400

Query Match 56.3%; Score 1282; DB 4; Length 346;
Best Local Similarity 62.5%; Pred. No. 1.2e-112;
Matches 262; Conservative 25; Mismatches 46; Indels 86; Gaps 3;
28 MVVKVGINGFGRIQRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRDGTVEV 87
14 MTVKVGINGFGRIQRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRDGTVEV 73
88 KGGFEVNGNFIKVSARDPENIDWATDGVLEALGTVVEVKGGFDVNGKFIKVSAAE 147

74 HEGSFVNGKEIKVLARNPEELPW----- 98
148 KPEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAREPANIDWATDGVLEIVLE 207
99 -----GEL----- 108
208 ATSFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTGTTETVISGASCTTN 267
109 CTGFFTSKKEAAEKHLTA-GAKRVVISAPGGNDVPTIVNTNHTLTGEEETVISGASCTTN 167
268 CLAPMAKALHDAPFQIGKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSGAA 327
168 CLAPMAKALHDNFGVWEGGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSGAA 227
328 KAIGLVIPELNGKLDGAAQRPVPTGTVTELVTLDKNVSVDEINAAKASNDSPGYTE 387
228 KAIGLVIPELNGKLDGAAQRPVPTGTVTELVTLDKNVSVDEINAAKASNDSPGYTE 287
388 DPTVSSDIVGVSGLSFDATQTKVMEVDGSQLKVVSWYDNEMSYTAQLVRLTLEYFAK 446
288 DEIVSSDIVGVSGLSFDATQTKVMEVDGSQLKVVSWYDNEMSYTAQLVRLTLEYFAK 346
RESULT 12
US-09-107-532A-4769
Sequence 4769, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4769:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...333
SEQUENCE DESCRIPTION: SEQ ID NO: 4769:
US-09-107-532A-4769

```

Query Match      55.8%; Score 1270; DB 4; Length 333;
Best Local Similarity 62.1%; Pred. No. 1.5e-111;
Matches 260; Conservative 26; Mismatches 47; Indels 86; Gaps 3;

QY 28 MVVKVINGFGRIGRLAPRRIONVEGVTRINDLTDNMLAHLKYDITQGRFDGTVEV 87
DB 1 MTVKVINGFGRIGRLAPRRIONVEGVTRINDLTDNMLAHLKYDITQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALGTVGVKGGFVNGKFIKVSAR 147
DB 61 HEGSFVNGKEVKVLANRNPPELW-
QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
DB 86 -----GEL-----GVDIVILE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTETVISGASCTTN 267
DB 96 CTGFTTSAAEKHLTA-GAKRVVISAFGNDVPIVNTNHEITLTKETVISGASCTTN 154
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
DB 155 CLAPMAKTLNDKFGVVEGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 214
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSVDEINAAKKAASNDSPGYTE 387
DB 215 KAIGLVIPELNGKYGAAQRPVPTGSGVTELVTLEKEVTVDENAAKKEASNESYGYNT 274
QY 388 DPVSSDVIWVSGYSLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKI 446
DB 275 DEIVSSDIVGMTYSGSLFDATQTKWTVGDKQLVKTVAWYDNEMSYTAQLVRTLEYFANL 333

RESULT 13
US-09-134-001C-5513
; Sequence 5513, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5513
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5513

Query Match      50.0%; Score 1138.5; DB 4; Length 348;
Best Local Similarity 55.7%; Pred. No. 4.2e-99;
Matches 235; Conservative 39; Mismatches 61; Indels 87; Gaps 4;

QY 28 MVVKVINGFGRIGRLAPRRIONVEGVTRINDLTDNMLAHLKYDITQGRFDGTVEV 87
DB 13 MAIKVAINGFGRIGRLAPRRIONVEGVTRINDLTDNMLAHLKYDITQGRFDGTVEV 72
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALGTVGVKGGFVNGKFIKVSAR 147
DB 73 IEGGFVNGKEIKSPDEPAGKLPW-----GDL----- 101
QY 148 KOPEQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
DB 102 -----IDVILE 107
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTETVISGASCTTN 267

```

```

DB 108 CTGFTTDEKAQAHAIDA-GAKKVLISAPAKGDKVKTIVFNTNHDILDTGSETVSGASCTTN 166
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
DB 167 SLAPVAKVLSDEFGLVEGFMTHAYTGDQNTQDAPHRKGDGRRARAAAENIIPNSTGAA 226
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSVDEINAAKKAASNDSPGYT 386
DB 227 KAIGKVIPEIDKLDGGAQRVPVATGSLTELTVVLDKQDVTVDQVNSAMKQASDESFGYT 286
QY 387 EDPVSSDIVGMTYSGYSLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKI 446
DB 287 EDEIVSSDIVGMTYSGSLFDATQTRVMTVGDRQLVKVAAWYDNEMSYTAQLVRTLAHLAEL 346
QY 447 AK 448
DB 347 SK 348

RESULT 14
US-09-134-000C-4229
; Sequence 4229, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4229
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4229

Query Match      40.5%; Score 923.5; DB 4; Length 357;
Best Local Similarity 47.2%; Pred. No. 8.5e-79;
Matches 199; Conservative 44; Mismatches 90; Indels 89; Gaps 6;

QY 28 MVVKVINGFGRIGRLAPRRIONV-EGVEVTRINDLTDNMLAHLKYDITQGRFDGTVE 86
DB 22 MTVKVINGFGRIGRLAPRRIOKVEVSDIEVVAINDLTSPTMLAHLQFDSHTGTYPGT 81
QY 87 VKEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALGTVGVKGGFVNGKFIKVSAR 146
DB 82 ATENGIVVDGGEVTRVYAPDASKIPW----- 107
QY 147 EKDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEV 206
DB 108 -----VKE-----NGVDIVL 117
QY 207 EATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTETVISGASCTT 266
DB 118 ECTGFTTSEKAQAHAIDA-GVKRVVISAPAG-AMKTIIVNVNDDTLTDNDKIIISGCTT 175
QY 267 NCLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGA 326
DB 176 NCLAPMAYFLANEFIEVGTMTTHAYTSTQMLDGPVRGNGNLRARSAANTIPHSSTA 235
QY 327 AKAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSVDEINAAKKAASNDSPG 384
DB 236 AKAIGLVIPELNGKLGQAQRVPVVDGSLTELVSILTKVTADQVNEAMKKHTIDNPSEG 295
QY 385 YTEDPIVSSDIVGYSGLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFA 444
DB 296 YDREIVSGDIIGTTEGSIPTQTEVTTAGDFQLVKTVAWYDNEYGTTCQIRLEKFA 355

```

QY 445 KI 446
Db 356 NL 357

RESULT 15

US-09-634-238-233
; Sequence 233, Application US/09634238
; Patent No. 654472
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christenson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 213
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-233

Query Match 40.3%; Score 918.5; DB 4; Length 340;
Best Local Similarity 46.0%; Pred. No. 2.3e-78;

Matches 196; Conservative 49; Mismatches 88; Indels 93; Gaps 7;

QY 28 MVKVGINGFRIGRIAFRIQNV---EGVEVTRINDLPNMLAHLKYDTTQGRFDG 83
Db 1 MTVKIGINGFRIGRIAFRIEIGAKSNDIQVVAINDLTSPTMLAHLKYDSTHGTFPG 60
QY 84 TVEVKEGGEVNGNFVKVSAERDPENIDWATDGEIVLEALEGTVEVKDGGFVNGKFTK 143
Db 61 EVSATDNGIVVDGKEYRVVAEPQAQNPW-----VKN----- 92
QY 144 VSAEKDPEQIDWATDGEIVLEIDGTVEVKEGGEVNGQFVKVSAEREPAIDWATDGE 203
Db 93 -----DGVD 96
QY 204 IVLEATSPFAKAEKHLHANGAKKVVITAPGNDVKTWFTNHDILDGTETVISGAS 263
Db 97 YVLECTGFYTSAEKSAHLDA-GAKVLSAPAGK-IKTIYVNVNDDTLNADDKIVSAGS 154
QY 264 CTTNCLAPNAKALHDAFGIOKGLMTTIHAYTGDQMLDGFHGGDLRRAGAANIVPNS 323
Db 155 CTTNCLAPMAYFLNCEFGIEVGTMTTHAYTSTQMLLDGPRVGNLRAARSAAANTIPHS 214
QY 324 TGAAGAIGLIVIPELNGKLDGAQRVPVPTGVTSELVVTI-DKNVSVDEINAMK--AASN 380
Db 215 TGAAGAIGLIVIPELNGKLGCHQRVSVVDGSGTSELVILKTRNVTADQVNEAIKHTENN 274
QY 381 DSFGYTEPIVSSDIVGVSGLSFATQTKVMEVDGSLVKVVSVDNEMSYTAQLVRTL 440
Db 275 PSFGWNEDEIVSSDIVGTYGSIFFDPTQTEVITAGDYQLVKVWYDNEYGFTCMIRTL 334
QY 441 EYFAKI 446
Db 335 LKFATL 340

Search completed: September 15, 2004, 10:11:44
Job time : 36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:08:58 ; Search time 131 Seconds
(without alignments)
1096.710 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278

Sequence: 1 MKKITGIILLAVILLSAC.....EMSYTAQLVRLTFYPAKIAK 448

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2278	100.0	448	9	US-09-878-766A-22
2	2278	100.0	448	12	US-10-650-369-22
3	1656.5	72.7	336	9	US-09-878-766A-12
4	1656.5	72.7	336	10	US-09-878-781-4
5	1656.5	72.7	336	12	US-10-650-369-12
6	1656.5	72.7	336	14	US-10-134-297-4
7	1655.5	72.7	336	12	US-10-282-122A-74379
8	1650	72.4	335	10	US-09-878-781-14
9	1562.5	68.6	336	10	US-09-878-781-16
10	1559.5	68.5	336	9	US-09-878-766A-16
11	1559.5	68.5	336	10	US-09-878-781-8
12	1559.5	68.5	336	12	US-10-650-369-16
13	1559.5	68.5	336	14	US-10-134-297-8
14	1557.5	68.4	336	9	US-09-878-766A-14
15	1557.5	68.4	336	10	US-09-878-781-6

16	1557.5	68.4	336	12	US-10-650-369-14	Sequence 14, Appl
17	1557.5	68.4	336	14	US-10-134-297-6	Sequence 6, Appl
18	1535	67.4	359	9	US-09-815-242-13169	Sequence 13169, A
19	1535	67.4	359	12	US-10-282-122A-74186	Sequence 74186, A
20	1534.5	67.4	336	9	US-09-878-766A-20	Sequence 20, Appl
21	1534.5	67.4	336	10	US-09-878-781-12	Sequence 12, Appl
22	1534.5	67.4	336	12	US-10-650-369-20	Sequence 20, Appl
23	1534.5	67.4	336	14	US-10-134-297-12	Sequence 12, Appl
24	1534	67.3	359	9	US-09-815-242-13593	Sequence 13593, A
25	1530.5	67.2	336	9	US-09-878-766A-18	Sequence 18, Appl
26	1530.5	67.2	336	10	US-09-878-781-10	Sequence 10, Appl
27	1530.5	67.2	336	12	US-10-650-369-18	Sequence 18, Appl
28	1530.5	67.2	336	14	US-10-134-297-10	Sequence 10, Appl
29	1521	66.8	337	12	US-10-282-122A-72254	Sequence 72254, A
30	1491	65.5	333	9	US-09-765-272-54	Sequence 54, Appl
31	1357.5	59.6	336	15	US-10-369-493-18524	Sequence 18524, A
32	1282	56.3	333	9	US-09-815-242-10847	Sequence 10847, A
33	1282	56.3	333	12	US-10-282-122A-57294	Sequence 57294, A
34	1280	56.2	337	15	US-10-369-493-18342	Sequence 18342, A
35	1224.5	53.8	335	12	US-10-282-122A-53266	Sequence 53266, A
36	1205.5	52.9	336	12	US-10-282-122A-60492	Sequence 60492, A
37	1203	52.8	334	12	US-10-282-122A-51519	Sequence 51519, A
38	1166	51.2	335	12	US-10-282-122A-71734	Sequence 71734, A
39	1145	50.3	334	12	US-10-282-122A-65661	Sequence 65661, A
40	1141	50.1	357	12	US-10-282-122A-65001	Sequence 65001, A
41	1138.5	50.0	336	12	US-10-282-122A-70890	Sequence 70890, A
42	1132.5	49.7	336	9	US-09-815-242-5618	Sequence 5618, Ap
43	1132.5	49.7	336	9	US-09-815-242-12540	Sequence 12540, A
44	1132.5	49.7	336	12	US-10-282-122A-43997	Sequence 43997, A
45	941.5	41.3	333	12	US-10-282-122A-56828	Sequence 56828, A

ALIGNMENTS

RESULT 1

```

US-09-878-766A-22
; Sequence 22, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878.766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
; OTHER INFORMATION: epitope fusion protein
US-09-878-766A-22

```

Query Match	100.0%	Score 2278;	DB 9;	Length 448;
Best Local Similarity	100.0%;	Pred. No. 1.2e-198;		
Matches 448;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MKKITGIILLAVILLSACQYGMVVKVINGFGRIGRLAFRRIONVEGVETRN	60	
Db	1	MKKITGIILLAVILLSACQYGMVVKVINGFGRIGRLAFRRIONVEGVETRN	60	
Qy	61	DLTPNMLHLKKYDTTQGRFDCGTVEVKEGSGFVNGNFIKUSAERDPENIDWATDGEIV	120	
Db	61	DLTPNMLHLKKYDTTQGRFDCGTVEVKEGSGFVNGNFIKUSAERDPENIDWATDGEIV	120	
Qy	121	LEALEGTVEVKGDFVNGKFIKUSAEDPQIDWATDGEIVLEIDGTVEVKEGSGV	180	

Db 121 LEAEGTVEVKGGFDVNGKFKVSAEKQPEQIDWATDGVIEIVLEIDGTVEVKEGGFEVN 180
 QY 181 GQFVKVSAEREPANIDWATDGVIEIVLEATSFPAKKEAAEKHLHANGAKKVVITAPGNDV 240
 Db 181 GQFVKVSAEREPANIDWATDGVIEIVLEATSFPAKKEAAEKHLHANGAKKVVITAPGNDV 240
 QY 241 KTVFNTNHDLDGTETVTSAGCTTNCCLAPMAKALHDAFGIQKGLMTTHAYTGDQMIL 300
 Db 241 KTVFNTNHDLDGTETVTSAGCTTNCCLAPMAKALHDAFGIQKGLMTTHAYTGDQMIL 300
 QY 301 DGHRRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTGVTSLV 360
 Db 301 DGHRRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTGVTSLV 360
 QY 361 TLDKNVSVDEINAAKAAASNDSTGYTDEPIVSSDIVGVSGLSFDATQTKWMEVDGSQLV 420
 Db 361 TLDKNVSVDEINAAKAAASNDSTGYTDEPIVSSDIVGVSGLSFDATQTKWMEVDGSQLV 420
 QY 421 KVSWSYDNEMSYTAQLVRTLEYFAKIAK 448
 Db 421 KVSWSYDNEMSYTAQLVRTLEYFAKIAK 448

RESULT 2

US-10-650-369-22

; Sequence 22, Application US/10650369

; Publication No. US20040062774A1

; GENERAL INFORMATION:

; APPLICANT: Potter, Andrew A.

; APPLICANT: Perez-Casal, Jose

; APPLICANT: Fontaine, Michael

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN

; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0057

; CURRENT APPLICATION NUMBER: US/10/650,369

; CURRENT FILING DATE: 2003-08-27

; PRIOR APPLICATION NUMBER: US/09/878,766A

; PRIOR FILING DATE: 2001-09-10

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 22

; LENGTH: 448

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple

; OTHER INFORMATION: epitope fusion protein

US-10-650-369-22

Query Match 100.0%; Score 2278; DB 12; Length 448;
 Best Local Similarity 100.0%; Pred. No. 1.2e-198;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKITGIILLAVIILSAQYAGVGVKGVNGFGRIGRLAFRRIONVEGVETRN 60
 Db 1 MKKITGIILLAVIILSAQYAGVGVKGVNGFGRIGRLAFRRIONVEGVETRN 60
 QY 61 DLTDPNMLAHLKYDTTQGRFDGTVEVKEGGFEVNGFVKVSAERDPENIDWATDGVIEIV 120
 Db 61 DLTDPNMLAHLKYDTTQGRFDGTVEVKEGGFEVNGFVKVSAERDPENIDWATDGVIEIV 120
 QY 121 LEAEGTVEVKGGFDVNGKFKVSAEKQPEQIDWATDGVIEIVLEIDGTVEVKEGGFEVN 180
 Db 121 LEAEGTVEVKGGFDVNGKFKVSAEKQPEQIDWATDGVIEIVLEIDGTVEVKEGGFEVN 180
 QY 181 GQFVKVSAEREPANIDWATDGVIEIVLEATSFPAKKEAAEKHLHANGAKKVVITAPGNDV 240
 Db 181 GQFVKVSAEREPANIDWATDGVIEIVLEATSFPAKKEAAEKHLHANGAKKVVITAPGNDV 240
 QY 241 KTVFNTNHDLDGTETVTSAGCTTNCCLAPMAKALHDAFGIQKGLMTTHAYTGDQMIL 300
 Db 241 KTVFNTNHDLDGTETVTSAGCTTNCCLAPMAKALHDAFGIQKGLMTTHAYTGDQMIL 300

QY 301 DGHRRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTGVTSLV 360
 Db 301 DGHRRGDLRRARAGAAIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTGVTSLV 360
 QY 361 TLDKNVSVDEINAAKAAASNDSTGYTDEPIVSSDIVGVSGLSFDATQTKWMEVDGSQLV 420
 Db 361 TLDKNVSVDEINAAKAAASNDSTGYTDEPIVSSDIVGVSGLSFDATQTKWMEVDGSQLV 420
 QY 421 KVSWSYDNEMSYTAQLVRTLEYFAKIAK 448
 Db 421 KVSWSYDNEMSYTAQLVRTLEYFAKIAK 448

RESULT 3

US-09-878-766A-12

; Sequence 12, Application US/09878766A

; Patent No. US20020044928A1

; GENERAL INFORMATION:

; APPLICANT: Potter, Andrew A.

; APPLICANT: Perez-Casal, Jose

; APPLICANT: Fontaine, Michael

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN

; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0057

; CURRENT APPLICATION NUMBER: US/09/878,766A

; CURRENT FILING DATE: 2001-09-10

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 12

; LENGTH: 336

; TYPE: PRT

; ORGANISM: Streptococcus dysgalactiae

US-09-878-766A-12

Query Match 72.7%; Score 1656.5; DB 9; Length 336;
 Best Local Similarity 79.6%; Pred. No. 3.4e-142;
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;
 QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRNLDTPNMLAHLKYDTTQGRFDGTVEV 87
 Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRNLDTPNMLAHLKYDTTQGRFDGTVEV 60
 QY 88 KEGGFVNGNFVKVSAERDPENIDWATDGVIEIVLEAEGTVEVKDGGFDVNGKFKVSAE 147
 Db 61 KEGGFVNGNFVKVSAERDPE----- 81
 QY 148 KQPEQIDWATDGVIEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEIV 207
 Db 82 -----NIDWATDGVIEIVLE 95
 QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVVNTNHDLDGTETVTSAGCTTN 267
 Db 96 ATGFFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVVNTNHDLDGTETVTSAGCTTN 155
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMILDGHRRGDLRRARAGAAIVPNSTGAA 327
 Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMILDGHRRGDLRRARAGAAIVPNSTGAA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVVTLDKNVSVDEINAAKAAASNDSTGYTE 387
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVVTLDKNVSVDEINAAKAAASNDSTGYTE 275
 QY 388 DPTVSSDIVGVSGLSFDATQTKWMEVDGSQLVKVSVSYDNEMSYTAQLVRTLEYFAKIA 447
 Db 276 DPTVSSDIVGVSGLSFDATQTKWMEVDGSQLVKVSVSYDNEMSYTAQLVRTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 4

US-09-878-781-4

```
; Sequence 4, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-09-878-781-4

Query Match      72.7%; Score 1656.5; DB 10; Length 336;
Best Local Similarity 79.6%; Pred. No. 3.4e-142;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

Qy 28 M V K V G I N G F G R I G L A F R R I Q N V E G V E T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 87
Db 1 M V K V G I N G F G R I G L A F R R I Q N V E G V E T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 60

Qy 88 K E G G F E V N G N F I K V S A E R D P E N I D W A T D G V E I V L E A L E G T V E V K D G G F V N G K F I K V S A E 147
Db 61 K E G G F E V N G N F I K V S A E R D P E ----- 81

Qy 148 K D P E I D W A T D G V E I V L E I D G T V E V K E G G F E V N G Q F V K V S A E R E P A N I D W A T D G V E I V L E 207
Db 82 ----- N I D W A T D G V E I V L E 95

Qy 208 A T S P F A K K E A E K H L H A N G A K K V I T A P G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 267
Db 96 A T G F F A K K E A E K H L H A N G A K K V I T A P G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 155

Qy 268 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D M I L D G P H R G G D L R R A G A A N I V P N S T G A A 327
Db 156 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D M I L D G P H R G G D L R R A G A A N I V P N S T G A A 215

Qy 328 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 387
Db 216 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 275

Qy 388 D P I V S S D I V G S Y S G S L F D A T Q T K M E V D G S Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 447
Db 276 D P I V S S D I V G S Y S G S L F D A T Q T K M E V D G S Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 335

Qy 448 K 448
Db 336 K 336

RESULT 5
US-10-650-369-12
; Sequence 12, Application US/10650369
; Publication No. US20040062774A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/10/650,369
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/878,766A
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22

Query Match      72.7%; Score 1656.5; DB 14; Length 336;
Best Local Similarity 79.6%; Pred. No. 3.4e-142;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

Qy 28 M V K V G I N G F G R I G L A F R R I Q N V E G V E T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 87
Db 1 M V K V G I N G F G R I G L A F R R I Q N V E G V E T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 60

Qy 88 K E G G F E V N G N F I K V S A E R D P E N I D W A T D G V E I V L E A L E G T V E V K D G G F V N G K F I K V S A E 147
Db 61 K E G G F E V N G N F I K V S A E R D P E ----- 81

Qy 148 K D P E I D W A T D G V E I V L E I D G T V E V K E G G F E V N G Q F V K V S A E R E P A N I D W A T D G V E I V L E 207
Db 82 ----- N I D W A T D G V E I V L E 95

Qy 208 A T S P F A K K E A E K H L H A N G A K K V I T A P G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 267
Db 96 A T G F F A K K E A E K H L H A N G A K K V I T A P G N D V K T V V F N T N H D I L D G T E T V I S G A S C T T N 155

Qy 268 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D M I L D G P H R G G D L R R A G A A N I V P N S T G A A 327
Db 156 C L A P M A K A L H D A F G I Q K G L M T T I H A Y T G D M I L D G P H R G G D L R R A G A A N I V P N S T G A A 215

Qy 328 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 387
Db 216 K A I G L V I P E L N G K L D G A A Q R V P V P T G S V T E L V T L D K N V S V D E I N A A M K A A S N D S F G Y T E 275

Qy 388 D P I V S S D I V G S Y S G S L F D A T Q T K M E V D G S Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 447
Db 276 D P I V S S D I V G S Y S G S L F D A T Q T K M E V D G S Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A 335

Qy 448 K 448
Db 336 K 336

RESULT 6
US-10-134-297-4
; Sequence 4, Application US/10134297
; Publication No. US20030165524A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055.20
; CURRENT APPLICATION NUMBER: US/10/134,297
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-10-134-297-4

Query Match      72.7%; Score 1656.5; DB 14; Length 336;
Best Local Similarity 79.6%; Pred. No. 3.4e-142;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

Qy 28 M V K V G I N G F G R I G L A F R R I Q N V E G V E T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 87
Db 1 M V K V G I N G F G R I G L A F R R I Q N V E G V E T R I N D L T D P N M L A H L L K Y D T T Q G R F D G T V E V 60
```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74379
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74379

Query Match      72.7%; Score 1655.5; DB 12; Length 336;
Best local Similarity 79.3%; Pred.No. 4.1e-142;
Matches 334; Conservative 1; Mismatches 1; Indels 85; Gaps 1;

Qy      28  MVVKVINGFGRIGRLAFRIQNVGVEVTRINDLTDPNMLAHLKKYDTTQGRFDGVVEV  87
      |||
Db      1  MVVKVINGFGRIGRLAFRIQNIIEGVETRIINDLTDPNMLAHLKKYDTTQGRFDGVVEV  60

Qy      88  KEGGFVNGNFIKVSARDPENIDWATGDVEIVLEALEGTVEVKDGFVNGRFIKVSAE  147
      |||
Db      61  KEGGFVNGNFIKVSARDPE-----  81

Qy      148  KDEPEQIDWATGDVEIVLEIDGTVEVKGGFEVNGQFVKVSAEREPANIDWATGVEIVLE  207
      |||
Db      82  -----NIDWATGVEIVLE  95

Qy      208  ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN  267
      |||
Db      96  ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVISGASCTTN  155

Qy      268  CLAPMAKALHDAFGIOKGLMTTHATVGQOMILDGPBGDLRRARAGAAIIVFNSTGAA  327
      |||
Db      156  CLAPMAKALHDAFGIOKGLMTTHATVGQOMILDGPBGDLRRARAGAAIIVFNSTGAA  215

Qy      328  KATGLVIPLENGKLDGAAQRPVPTGSVTLELVTLDKNVSVDSVEINAAKKAASNDSFGYTE  387
      |||
Db      216  KATGLVIPLENGKLDGAAQRPVPTGSVTLELVTLDKNVSVDSVEINAAKKAASNDSFGYTE  275

Qy      388  DPVSSDIYGVSYGSLFDATQTKMVEVDGSQLVKVYVDNENMSYTAQLVRTLEYFAKIA  447
      |||
Db      276  DPVSSDIYGVSYGSLFDATQTKMVEVDGSQLVKVYVDNENMSYTAQLVRTLEYFAKIA  335

Qy      448  K 448
      |
Db      336  K 336

```

```

RESULT 8
US-09-878-781-14
; Sequence 14, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; FILE REFERENCE: STREPTOCOCCUS INFECTION
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SpyGapC
; OTHER INFORMATION: protein
US-09-878-781-14

Query Match              72.4%; Score 1650; DB 10; Length 335;
Best Local Similarity    79.1%; Pred. No. 1.3e-141;
Matches 333; Conservative 2; Mismatches 0; Indels 86; Gaps 1;

```



```

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNNLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNNLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFVNGKFIKVSAAE 147
Db 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 82 -----IDWATDGVIEVLE 94
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 95 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 154
QY 268 CLAPMAKALHDFAFGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAANIIPVNSTGAA 327
Db 155 CLAPMAKALHDFAFGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAANIIPVNSTGAA 214
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLTKNVSVDEINAAKKAASNDSPGYTE 387
Db 215 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLTKNVSVDEINAAKKAASNDSPGYTE 274
QY 388 DPIVSSDIVGVSGLFDATQTKMVEVDSQIVKVVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 275 DPIVSSDIVGVSGLFDATQTKMVEVDSQIVKVVSWYDNEMSYTAQLVRLTLEYFAKIA 334
QY 448 K 448
Db 335 K 335
RESULT 9
US-09-878-781-16
; Sequence 16, Application US/09878781
; Patent No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; FILE REFERENCE: 9000-0055
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SeqGapC
; OTHER INFORMATION: protein
US-09-878-781-16
Query Match 68.6%; Score 1562.5; DB 10; Length 336;
Best Local Similarity 76.2%; Pred. No. 1.2e-133;
Matches 321; Conservative 2; Mismatches 13; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNNLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNNLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGGFVNGKFIKVSAAE 147
Db 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 82 -----NIDWATDGVIEVLE 95

```

```

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 96 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDFAFGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAANIIPVNSTGAA 327
Db 156 CLAPMAKALHDFAFGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAANIIPVNSTGAR 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLTKNVSVDEINAAKKAASNDSPGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLTKNVSVDEINAAKKAASNDSPGYTE 275
QY 388 DPIVSSDIVGVSGLFDATQTKMVEVDSQIVKVVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPIVSSDIVGVSGLFDATQTKMVEVDSQIVKVVSWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336
RESULT 10
US-09-878-766A-16
; Sequence 16, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-766A-16
Query Match 68.5%; Score 1559.5; DB 9; Length 336;
Best Local Similarity 73.9%; Pred. No. 2.3e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNNLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNNLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALSGTVEVKDGGFVNGKFIKVSAAE 147
Db 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 82 -----NIDWATDGVIEVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 96 ATSPFAKKEAAEKHLHANGAKKVVITAPGGDDVKTVVFNTHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDFAFGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAANIIPVNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDQMIIDGPHRGDLRRARAGASNIIPVNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLTKNVSVDEINAAKKAASNDSPGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLTKNVSVDEINAAKKAASNDSPGYTE 275
QY 388 DPIVSSDIVGVSGLFDATQTKMVEVDSQIVKVVSWYDNEMSYTAQLVRLTLEYFAKIA 447

```

Db 276 DPVSSDIIGMAYGSLFDATQTKVQTVGDNQLVKVWSWYDNEMSYTAQLVRTLEYFAKIA 335
Qy 448 K 448
Db 336 K 336

RESULT 11

US-09-878-781-8
; Sequence 8, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; FILE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-781-8

Query Match 68.5%; Score 1559.5; DB 10; Length 336;
Best Local Similarity 73.9%; Pred. No. 2.3e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
Qy 28 MVMKVGINGFGRIGRLAFRRIQNVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVMKVGINGFGRIGRLAFRRIQNVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Qy 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVKDGDFVNGKFIKVSAAE 147
Db 61 KGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVKDGDFVNGKFIKVSAAE 81
Qy 148 KDEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLE 207
Db 82 -----NIDWATDGVLE 95
Qy 208 ATSPFAKKEAARKHLHANGAKKVVITAPGNDVKTWVNTNHDILDTETVTSASCTTN 267
Db 96 ATGFFAKKAAAEKHLHANGAKKVVITAPGDDVKTWVNTNHDILDTETVTSASCTTN 155
Qy 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDPHRRGGDLRRARAGASNIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVKGQGLMTTHAYTGDMILDPHRRGGDLRRARAGASNIVPNSTGAA 215
Qy 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLDKNVSDVEINAAKAAASNDGFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLVLEKETSVEEINAAKAAASNDGFGYTE 275
Qy 388 DPVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
Db 276 DPVSSDIIGMAYGSLFDATQTKVQTVGDNQLVKVWSWYDNEMSYTAQLVRTLEYFAKIA 335
Qy 448 K 448
Db 336 K 336

RESULT 12

US-10-650-369-16
; Sequence 16, Application US/10650369
; Publication No. US20040062774A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose

; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/10/650,369
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/878,766A
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-10-650-369-16

Query Match 68.5%; Score 1559.5; DB 12; Length 336;
Best Local Similarity 73.9%; Pred. No. 2.3e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
Qy 28 MVMKVGINGFGRIGRLAFRRIQNVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVMKVGINGFGRIGRLAFRRIQNVEGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
Qy 88 KEGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVKDGDFVNGKFIKVSAAE 147
Db 61 KGGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVKDGDFVNGKFIKVSAAE 81
Qy 148 KDEQIDWATDGVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVLE 207
Db 82 -----NIDWATDGVLE 95
Qy 208 ATSPFAKKEAARKHLHANGAKKVVITAPGNDVKTWVNTNHDILDTETVTSASCTTN 267
Db 96 ATGFFAKKAAAEKHLHANGAKKVVITAPGDDVKTWVNTNHDILDTETVTSASCTTN 155
Qy 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDPHRRGGDLRRARAGASNIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVKGQGLMTTHAYTGDMILDPHRRGGDLRRARAGASNIVPNSTGAA 215
Qy 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLDKNVSDVEINAAKAAASNDGFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLVLEKETSVEEINAAKAAASNDGFGYTE 275
Qy 388 DPVSSDIVGSYGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
Db 276 DPVSSDIIGMAYGSLFDATQTKVQTVGDNQLVKVWSWYDNEMSYTAQLVRTLEYFAKIA 335
Qy 448 K 448
Db 336 K 336

RESULT 13

US-10-134-297-8
; Sequence 8, Application US/10134297
; Publication No. US20030165524A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; FILE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055.20
; CURRENT APPLICATION NUMBER: US/10/134,297
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis

US-10-134-297-8

```
Query Match      68.5%; Score 1559.5; DB 14; Length 336;
Best Local Similarity 73.9%; Pred. No. 2.3e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRFDGTVEV 87
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVBIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
DB 61 KGGFVNGNFIKVSARDPE----- 81
QY 148 KDPEQIDWATDGVBIVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVBIVLE 207
DB 82 -----NIDWATDGVBIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 267
DB 96 ATGFFAKKAAEKHLHANGAKKVVITAPGGDDVKTVVNTNHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
DB 156 CLAPMAKALQDNFVKGQGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSVDEINAAKAAANDSFGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVALEKDVTVVEVNAAMKAAANDSFGYTE 275
QY 388 DPIVSSDIVGSYGLFDATQTKVMEVDGSQLVKVWSWYDNEMSYTAQLVRLTLEYFAKIA 447
DB 276 DPIVSSDIVGSYGLFDATQTKVQTVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 335
QY 448 K 448
DB 336 K 336

RESULT 14
US-09-878-766A-14
; Sequence 14, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-766A-14

Query Match      68.4%; Score 1557.5; DB 9; Length 336;
Best Local Similarity 73.9%; Pred. No. 3.5e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRFDGTVEV 87
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRF----- 54
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVBIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
DB 55 ----- 54
QY 148 KDPEQIDWATDGVBIVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVBIVLE 207
DB -----DGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVBIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 267
DB 96 ATGFFAKKAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
DB 156 CLAPMAKALQDNFVKGQGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSVDEINAAKAAANDSFGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVALEKDVTVVEVNAAMKAAANDSFGYTE 275

Query Match      68.4%; Score 1557.5; DB 10; Length 336;
Best Local Similarity 73.9%; Pred. No. 3.5e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRFDGTVEV 87
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTDPNMLAHLKYDTTQGRF----- 54
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVBIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
DB 55 ----- 54
QY 148 KDPEQIDWATDGVBIVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVBIVLE 207
DB -----DGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVBIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 267
DB 96 ATGFFAKKAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
DB 156 CLAPMAKALQDNFVKGQGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSVDEINAAKAAANDSFGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVALEKDVTVVEVNAAMKAAANDSFGYTE 275
```

QY 388 DPIVSSDIVGVSIGSLFDATOTKMEVDGSQLVKVSWYDNMSYTAQLVRTILEYFAKIA 447
Db 276 DPIVSSDIVGVSIGSLFDATOTKMEVDGSQLVKVSWYDNMSYTAQLVRTILEYFAKIA 335

QY 448 K 448
Db 336 K 336

Search completed: September 15, 2004, 10:22:15
Job time : 133 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:20:04 ; Search time 41 Seconds

(without alignments)
1051.068 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278

Sequence: 1 MKKITGIILLLLAVILSAC.....EMSYAQLVTLTLEFAKIAK 448

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1652.5	72.5	336	2	A42963	glyceraldehyde-3-p
2	1558.5	68.4	336	2	S71350	glyceraldehyde-3-p
3	1535	67.4	335	2	F95235	glyceraldehyde 3-p
4	1535	67.4	359	2	G98099	glyceraldehyde-3-p
5	1357.5	59.6	336	2	F86905	hypothetical prote
6	1280	56.2	337	2	G86694	hypothetical prote
7	1234	54.2	334	2	S34254	glyceraldehyde-3-p
8	1205.5	52.9	336	2	AC1382	glyceraldehyde 3-p
9	1204.5	52.9	336	2	AD1751	glyceraldehyde 3-p
10	1203	52.8	334	2	C96987	glyceraldehyde-3-p
11	1145	50.3	334	2	B82019	glyceraldehyde-3-p
12	1132.5	49.7	336	2	E89850	glyceraldehyde 3-p
13	1132	49.7	334	2	E81001	glyceraldehyde-3-p
14	955.5	41.9	338	2	T09633	glyceraldehyde 3-p
15	941.5	41.3	333	2	F90881	glyceraldehyde-3-p
16	941.5	41.3	333	2	C85737	glyceraldehyde-3-p
17	870.5	38.2	337	2	S73737	glyceraldehyde-3-p
18	861.5	37.8	349	2	F90517	glyceraldehyde 3-p
19	854.5	37.5	337	2	C64233	glyceraldehyde-3-p
20	851	37.4	336	2	A43260	glyceraldehyde-3-p
21	849	37.3	335	1	DEBSG	glyceraldehyde-3-p
22	845	37.1	335	2	S12696	glyceraldehyde-3-p
23	836	36.7	335	2	H84094	glyceraldehyde-3-p
24	830	36.4	335	1	DEBSGF	glyceraldehyde-3-p
25	820.5	36.0	335	2	A70107	probable glycerald
26	819.5	36.0	333	1	DEHGGT	glyceraldehyde-3-p
27	793.5	34.8	342	2	F70391	glyceraldehyde-3-p
28	790	34.7	336	2	T36020	glyceraldehyde-3-p
29	786.5	34.5	339	2	G70915	glyceraldehyde-3-p

ALIGNMENTS

RESULT 1

A42963
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus
N;Alternate names: plasmin receptor
C;Species: Streptococcus sp.
C;Date: 10-Mar-1994 #sequence revision 10-Mar-1994 #text_change 03-Jun-2002
C;Accession: A42963; B42963; JH0750
R;Lottenberg, R.; Broder, C.C.; Boyle, M.D.; Kain, S.J.; Schroeder, B.L.; Curtiss III, R.
J. Bacteriol. 174, 5204-5210, 1992
A;Title: Cloning, sequence analysis, and expression in Escherichia coli of a streptococcal
A;Reference number: A42963; MUID:92355491; PMID:1322883
A;Accession: A42963
A;Molecule type: DNA
A;Residues: 1-336 <LOT>
A;Experimental source: Group A, strain 64/14
A;Note: sequence extracted from NCBI backbone (NCBIP:110308)
A;Accession: B42963
A;Molecule type: protein
A;Residues: 2-74;161-164, 'X',166-174;187-211, 'X',213-217 <LO2>
R;Pancholi, V.; Fischetti, V.A.
J. Exp. Med. 176, 415-426, 1992
A;Title: A major surface protein on group A streptococci is a glyceraldehyde-3-phosphate
A;Reference number: JH0750; MUID:92364544; PMID:1500854
A;Accession: JH0750
A;Molecule type: protein
A;Residues: 2-30, 'A',32-40 <PAN>
C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C;Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
F;152/Active site: Cys #status predicted

Query Match 72.5%; Score 1652.5; DB 2; Length 336;
Best Local Similarity 79.1%; Pred. No. 9.1e-101;
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;

QY	28	MVVKVINGFGRIGRLAFRRIONVEGVETVRINDLTPNMLAHLLKYDTTQGRFDGTVEV	87
DB	1	MVVKVINGFGRIGRLAFRRIONVEGVETVRINDLTPNMLAHLLKYDTTQGRFDGTVEV	60
QY	88	KEGGFEVNGNFIKVSARPDENIDWATDGVETVLEALEGTVKDGDFVNGFIKVSAR	147
DB	61	KEGGFEVNGNFIKVSARPDENIDWATDGVETVLEALEGTVKDGDFVNGFIKVSAR	81
QY	148	KDPEQIDWATDGVETVLEIDGTVEVKEGGFEVNGFVKVSAEREPANIDWATDGVETVLE	207
DB	82	-----NIDWATDGVETVLE	95
QY	208	ATSFPAKKEAAEKHLHANGAKKVVITAPGGNDVKVTVPNTNHDILDTETVIGASCTTN	267
DB	96	ATGFFPAKKEAAEKHLHANGAKKVVITAPGGNDVKVTVPNTNHDILDTETVIGASCTTN	155
QY	268	CLAPMAKALHDAFGIOKGLMTTHAYTGDMIDGPHRGDLRRARAGANIVPNSGAA	327

Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAIVPNSGTAA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLIDKNVSDVEINAAKKAASNDGFGYTE 387
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLIDKNVSDVEINAAKKAASNDGFGYTE 275
 QY 388 DPVSSDIVGVSYSGLFDATQTKMEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA 447
 Db 276 DPVSSDIVGVSYSGLFDATQTKMEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 2

S71350
 C:Species: Streptococcus "equisimilis"
 C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 03-Jun-2002
 C:Accession: S71350
 R:Gase, K.; Gase, A.; Schirmer, H.; Malke, H.
 Eur. J. Biochem. 239, 42-51, 1996
 A:Title: Cloning, sequencing and functional overexpression of the Streptococcus equisimilis protein. Purification and biochemical characterization of the protein.
 A:Reference number: S71350; MUID:96305364; PMID:8706717
 A:Accession: S71350
 A:Molecule type: DNA
 A:Residues: 1-336 <GAS>
 A:Cross-references: EMBL:X97789; NID:g1478268; PIDN:CAA66377.1; PID:g1478269
 A:Experimental source: strain H46A
 C:Genetics:
 A:Gene: gapC

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
 C:Keywords: Gluconeogenesis; glycolysis; NAD; oxidoreductase
 F:4-34/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:152/Active site: Cys #status predicted

Query Match 68.4%; Score 1558.5; DB 2; Length 336;
 Best Local Similarity 76.0%; Pred. No. 1.3e-94;
 Matches 320; Conservative 2; Mismatches 14; Indels 85; Gaps 1;
 QY 28 MVKVGINGFGRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 87
 Db 1 MVKVGINGFGRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 60
 QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVKVGDFVNGFKIKVSAE 147
 Db 61 KEGGFVNGNFIKVSARDPE-----81
 QY 148 KOPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 Db 82 -----NIDWATDGVIEVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTTVTSIGASCTTN 267
 Db 96 ATGFFAKKEAAEKPLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTTVTSIGASCTTN 155
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAIVPNSGTAA 327
 Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAIVPNSGTAR 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLIDKNVSDVEINAAKKAASNDGFGYTE 387
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLIDKNVSDVEINAAKKAASNDGFGYTE 275
 QY 388 DPVSSDIVGVSYSGLFDATQTKMEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA 447
 Db 276 DPVSSDIVGVSYSGLFDATQTKMEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 3

F95235
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C:Accession: F95235
 R:Tetzelin, H.; Neilson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heaton, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, M.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: F95235
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-335 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK76079.1; PID:g14973522; GSPDB:GN00164; TIGR:S71350
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP2012
 C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 67.4%; Score 1535; DB 2; Length 335;
 Best Local Similarity 73.6%; Pred. No. 4.4e-93;
 Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
 QY 28 MVKVGINGFGRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 87
 Db 1 MVKVGINGFGRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRFDGTVVEV 60
 QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVKVGDFVNGFKIKVSAE 147
 Db 61 KEGGFVNGNFIKVSARDPE-----81
 QY 148 KOPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 Db 82 -----QIDWATDGVIEVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTTVTSIGASCTTN 267
 Db 96 ATGFFAKKEAAEKPLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTTVTSIGASCTTN 154
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAIVPNSGTAA 327
 Db 155 CLAPMAKALQDNFVGVVGLMTTHAYTGDQMLDGHRRGDLRRARAGAAIVPNSGTAA 214
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLIDKNVSDVEINAAKKAASNDGFGYTE 387
 Db 215 KAIGLVIPELNGKLDGSAQRVPTGVSVELVAVLEKNTVDENVNAAKAASNESYGYTE 274
 QY 388 DPVSSDIVGVSYSGLFDATQTKMEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA 447
 Db 275 DPVSSDIVGVSYSGLFDATQTKLVDPKQLVKVSWYDNMSYTAQLVRLTLEYFAKIA 334
 QY 448 K 448
 Db 335 K 335

RESULT 4

G98099
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 03-Jun-2002
 C:Accession: G98099
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Ede, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Myer, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A;Reference number: A97872; MUID:21429245; PMID:11544234

A;Accession: G98099

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-359 <KUR>

A;Cross-references: GB:AE007317; PIDN:AA00628.1; PID:g15459513; GSPDB:GN00174

C;Genetics:

A;Gene: gapA

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C;Keywords: oxidoreductase

Query Match 67.4%; Score 1535; DB 2; Length 359;
Best Local Similarity 73.6%; Pred. No. 4.8e-93;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVKVGINGFGRIGRLAFRR:QNVEGVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 87

DB 25 MVKVGINGFGRIGRLAFRR:QNVEGVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 84

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGEIVLEALEGTVKDGDFVNGKFIKVSAAE 147

DB 85 KEGGFVNGKFIKVSARDPE----- 105

QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207

DB 106 -----QIDWATDGEIVLE 119

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDTGTTETVISGASCTTN 267

DB 120 ATGFFAKKEAAEKHL-KGGAKKVVITAPGGNDVKTVFNTNHDVLDGTETVISGASCTTN 178

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAINVPNSTGAA 327

DB 179 CLAPMAKALQNFVGVVGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAINVPNSTGAA 238

QY 328 KAIGLVIPELNGKLDGAQRVPVPTGTVTLVVDLKNVSVDEINAAKKAASNDSPFGYTE 387

DB 239 KAIGLVIPELNGKLDGAQRVPVPTGTVTLVVDLKNVSVDEINAAKKAASNDSPFGYTE 298

QY 388 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447

DB 299 DPIVSSDIVGVSGLFDATQTKVLDVKGQKLVKVSVDNEMSYTAQLVRLTLEYFAKIA 358

QY 448 K 448

DB 359 K 359

RESULT 5

F86905

hypothetical protein gapB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C;Species: Lactococcus lactis subsp. lactis

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C;Accession: F86905

R;Polotin, A.; Wincker, P.; Mager, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss

A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: F86905

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-336 <STO>

A;Cross-references: GB:AE005176; PID:g12725315; PIDN:AAK06344.1; GSPDB:GN00146

A;Experimental source: strain IL1403

C;Genetics:

A;Gene: gapB

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 59.6%; Score 1357.5; DB 2; Length 336;
Best Local Similarity 65.1%; Pred. No. 1.8e-81;
Matches 274; Conservative 21; Mismatches 41; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRR:QNVEGVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 87

DB 1 MVKVGINGFGRIGRLAFRR:QNVEGVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 54

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGEIVLEALEGTVKDGDFVNGKFIKVSAAE 147

DB 55 ----- 54

QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207

DB 55 -----DGGKEVVDGGEVNGQFVKVTAESPANINWAEVGAIEVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDTGTTETVISGASCTTN 267

DB 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGSDVATVFNTHNVDLDTGTTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAINVPNSTGAA 327

DB 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAINVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAQRVPVPTGTVTLVVDLKNVSVDEINAAKKAASNDSPFGYTE 387

DB 216 KAIGLVIPELNGKLDGAQRVPVPTGTVTLVVDLKNVSVDEINAAKKAASNDSPFGYTE 275

QY 388 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447

DB 276 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448

DB 336 K 336

RESULT 6

G86694

hypothetical protein gapA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C;Species: Lactococcus lactis subsp. lactis

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C;Accession: G86694

R;Polotin, A.; Wincker, P.; Mager, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss

A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: G86694

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-337 <STO>

A;Cross-references: GB:AE005176; PID:g12723446; PIDN:AAK04657.1; GSPDB:GN00146

A;Experimental source: strain IL1403

C;Genetics:

A;Gene: gapA

C;Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 56.2%; Score 1280; DB 2; Length 337;
Best Local Similarity 61.4%; Pred. No. 2e-76;
Matches 259; Conservative 31; Mismatches 46; Indels 86; Gaps 2;

QY 28 MVKVGINGFGRIGRLAFRR:QNVEGVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 87

DB 1 MVKVGINGFGRIGRLAFRR:QNVEGVETRIINDLTDNMLAHLKYDTTQGRFDGTVEV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGEIVLEALEGTVKDGDFVNGKFIKVSAAE 147

DB 61 KEGGFVNGKFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207

DB 82 -----DIQWADSGVEIVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDTGTTETVISGASCTTN 267

DB 96 ATGFFATKEAEKHLHPGGAKKVITAPGGNDVKTVFNTNHDILDTGTTETVISGASCTTN 155

```
QY 268 CLAPMAKALHDAGFQIGKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAAIVPNSTGAA 327
Db 156 SLAPMADALNKNFVGKGTMTTSHSYTGDQMTLDGPHRGDFRARRAAAENIVPASSGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVTVLDKNVSVDEINAAKAAASNDSPGYTE 387
Db 216 KAIGLVLPELNGKLDGAAQRPVPTGVTSELVTVLDKNVSVDEINAAKAAASNDSPGYTV 275
QY 388 DPVSSDIVGVSGLSFLDQTKMVEV-DGSQLVKVSWYDNEMSYTAQLVRLTEYFAKI 446
Db 276 DEIVSSDLIGMAYGSLFDALTEVTDLKGQQLVKTAARYDNEMSFQAQLIRLTLEYFAKI 335
QY 447 AK 448
Db 336 AK 337

RESULT 7
S34254
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Clostridium p
C:Species: Clostridium pasteurianum
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 03-Jun-2002
C:Accession: S34254
R:Oster, T.; Asschoei, O.; Scherrer, S.; Branlant, G.; Branlant, C.
submitted to the EMBL Data Library, May 1993
A:Description: Nucleotide sequence of the glyceraldehyde-3-phosphate dehydrogenase from
A:Reference number: S34254
A:Accession: S34254
A:Molecule type: DNA
A:Residues: 1-334 <OST>
A:Cross-references: EMBL:X72219; NID:g311923; PIDN:CAA51020.1; PID:g311924
C:Superfamily: Glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: oxidoreductase

Query Match 54.2%; Score 1234; DB 2; Length 334;
Best Local Similarity 59.3%; Pred. No. 2.1e-73;
Matches 249; Conservative 34; Mismatches 51; Indels 86; Gaps 3;
QY 29 VVKVINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGVEVK 88
Db 1 MTKVAINGFGRIGRLARILEVPGLEVVAINDLTDKMLAHLFKYDSSQGRFNGIEVK 60
QY 89 EGGFEVGNFPIKVSARDPENIDWATDGVVEIVLEALEGTVEVKGDFVNGKFIKVSABK 148
Db 61 EGAFFVNGKEVKVFAADPEKLPW-----GEL-----GIDVLEK 95
QY 149 DPEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLEA 208
Db 85 -----GEL-----GIDVLEK 95
QY 209 TFFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDTETVVISGASCTTNC 268
Db 96 TGFFTKKEAAHVRA-GAKKVVISAPAGNDLKTIVNVNNEIDLDTETVVISGASCTTNC 154
QY 269 LAPMAKALHDAGFQIGKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAAIVPNSTGAAK 328
Db 155 LAPMAKVLNDFGIEKGFMTTHAYTNDQNTLDGPHRGDFRARRAAAENIVPNSTGAAK 214
QY 329 AIGLVIPELNGKLDGAAQRPVPTGVTSELVTVLDKNVSVDEINAAKAAASNDSPGYTED 388
Db 215 ATAQVLPELNGKLDGAAQRPVPTGVTSELVTVLDKNVSVDEINAAKAAASNDSPGYTED 274
QY 389 PIVSSDIVGVSGLSFLDQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 448
Db 275 EIVSADVVGISGLSFLDALTITKIVDVGSQLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 334

RESULT 8
AC1382
glyceraldehyde 3-phosphate dehydrogenase homolog gap [imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AC1382
```

```
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krest, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1382
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-336 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00537.1; PID:g16411947; GSPDB:GNC0177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: gap
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 52.9%; Score 1205.5; DB 2; Length 336;
Best Local Similarity 58.5%; Pred. No. 1.5e-71;
Matches 247; Conservative 34; Mismatches 54; Indels 87; Gaps 4;
QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGVEV 87
Db 1 MTKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGVEV 60
QY 88 KEGGFEVGNFPIKVSARDPENIDWATDGVVEIVLEALEGTVEVKGDFVNGKFIKVSABK 147
Db 61 HDGFFKVGKEVKVLANRNPPELPWG-----DL-----GVDIVLE 95
QY 148 KDEQIDWATDGVVEIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207
Db 89 -----GVDIVLE 95
QY 208 ATFFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVFNTNHDILDTETVVISGASCTTN 267
Db 96 CTGFFTAQDKAEHLHKA-GAKKVVISAPATGDMKTIVNVNNEIDLDTETVVISGASCTTN 154
QY 268 CLAPMAKALHDAGFQIGKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAAIVPNSTGAA 327
Db 155 CLAPMAKVLNDFGVEGVETRIINDLTPNMLAHLKYDTTQGRFDGVEV 214
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSELVTVLDKNVSVDEINAAKAAASNDSPGYT 386
Db 215 KAIGEVLPFLKGLDGAARVPVPTGVTSELVTVLDKNVSVDEINAAKAAASNDSPGYT 274
QY 387 EPIVSSDIVGVSGLSFLDQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTEYFAKI 446
Db 275 SDQVSSDIKGMTGSLFDQTKVLTGVDQQLVKTVAWYDNEMSYTAQLVRLTEYFAKI 334
QY 447 AK 448
Db 335 AK 336

RESULT 9
AD1751
glyceraldehyde 3-phosphate dehydrogenase homolog gap [imported] - Listeria innocua
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AD1751
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krest, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1751
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-336 <GLA>
```


A:Cross-references: GB:AL592022; PIDN:CAC97780.1; PID:gl6415075; GSPDB:GN00178
A:Experimental source: strain Clip1262

C:Genetics:

A:Gene: gap

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 52.9%; Score 1204.5; DB 2; Length 336;
Best Local Similarity 58.5%; Pred. No. 1.8e-71; Indels 87; Gaps 4;
Matches 247; Conservative 33; Mismatches 55;

QY 28 MVKVGINGFGRIGRLAFRRIQNVGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87

DB 1 MTVKVGINGFGRIGRLAFRRIQNVGIEVVAINDLTDKMLAHLKYDTTQGRFDGEV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATDQVEIVLEALRGTVVEVKDGGFDVNGKFIKVS 147

DB 61 HDGFVNGKEVKVLANRNPBLPWG-----DL----- 88

QY 148 KDPEQIDWATDQVEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDQVEIVLE 207

DB 89 -----GVDIVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTETVISGASCTTN 267

DB 96 CTGFTAQDKAELHIKA-GAKKVVISAPATGDMKTIYVNVNHEITLDGTETVISGASCTTN 154

QY 268 CLAPMAKALHDFAFGQKGLMTTIHAYTGDQMLDGPGRGDLRRRAGAANIVPNSTGAA 327

DB 155 CLAPMAKVLNDFKFGIEKGFMTTIHAYTGDQNTLDGPHRKGLRRAPAAAVIIPNSTGAA 214

QY 328 KAIGLVIPELNGKLDGAAQORVPVPTGSGVTVELVTLDKNVSVDEINAAKAAASN-DSFGYT 386

DB 215 KAIGEVLPFLKGLDGAAGAAQORVPVPTGSGVTVELVTLDKNVSVDEINAAKAAASDPFTFGT 274

QY 387 EDPIVSSDIVGVSGLFDATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKI 446

DB 275 SDQVSSDIKGMTFGLFDTQTKVLTGQDLVKTVAWYDNEMSYTAQLVRLTLEYFAKI 334

QY 447 AK 448

DB 335 AK 336

RESULT 10

C96987

glyceraldehyde 3-phosphate dehydrogenase, gene gapC [imported] - Clostridium acetobutyli

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C:Accession: C96987

R:Noelling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: C96987

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-334 <XUR>

A:Cross-references: GB:AE001437; PIDN:AAK78686.1; PID:gl5023589; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC0709

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 52.8%; Score 1203; DB 2; Length 334;

Best Local Similarity 57.4%; Pred. No. 2.2e-71;

Matches 241; Conservative 41; Mismatches 52; Indels 86; Gaps 3;

QY 29 VVKVGINGFGRIGRLAFRRIQNVGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 88

DB 1 MAKIAINGFGRIGRLALRILEVPGLEVVAINDLTDKMLAHLFKYDSSQGRFNGEIEVK 60

QY 89 EGGFVNGNFIKVSARDPENIDWATDQVEIVLEALRGTVVEVKDGGFDVNGKFIKVS 148

DB 61 EGAFVNGKEVKVFAEADPEKLPWG-----DL----- 87

QY 149 DPEQIDWATDQVEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDQVEIVLE 208

DB 88 -----GIDVLE 95

QY 209 TSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTETVISGASCTTN 268

DB 96 TGFTFKKEKAHAHRA-GAKKVVISAPAGNDKTIYVNVNNEEDLDGTETVISGASCTTN 154

QY 269 LAPMAKALHDFAFGQKGLMTTIHAYTGDQMLDGPGRGDLRRRAGAANIVPNSTGAA 328

DB 155 LAPMAKVLNDFKFGIEKGFMTTIHAYTGDQNTLDGPHRKGLRRAPAAAVIIPNSTGAA 214

QY 329 AIGLVIPELNGKLDGAAQORVPVPTGSGVTVELVTLDKNVSVDEINAAKAAASNDSEGYTED 388

DB 215 AISQVLPDLAKGLDGAQORVPVPTGSGVTVELVTLDKNVSVDEINAAKAAASNDSEGYTED 274

QY 389 PIVSSDIVGVSGLFDATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIAK 448

DB 275 PIVSADVVGINGVSLFDATLTKIVDVNGSQLVKTAAWYDNEMSYTSLVRLTLEYFAKIAK 334

RESULT 11

B32019

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) C NMA0246 [impo

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002

C:Accession: B32019

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: B32019

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-334 <PAR>

A:Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83554.1; PID:g737900

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA0246

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: oxidoreductase

Query Match 50.3%; Score 1145; DB 2; Length 334;

Best Local Similarity 56.4%; Pred. No. 1.4e-67;

Matches 235; Conservative 35; Mismatches 61; Indels 86; Gaps 4;

QY 28 MVKVGINGFGRIGRLAFRRIQNVGVEVTRINDLTDPNMLAHLKYDTTQGRFDGTVEV 87

DB 1 MSIKVAINGFGRIGRLALRLEKAGIEVVAINDLTPAEMLLHLFKYDSTQGRF----- 54

QY 88 KEGGFVNGNFIKVSARDPENIDWATDQVEIVLEALRGTVVEVKDGGFDVNGKFIKVS 147

DB 55 -----OQTAEKDDAIVNGKEIKVFAN 77

QY 148 KDPEQIDWATDQVEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDQVEIVLE 207

DB 78 PNPEELPW-----GEL-----GVDVVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTETVISGASCTTN 267

DB 96 CTGFTNTKAEAHIRA-GARKVVISAPGNDVKTVYGVNQDILDSSETVISASACTTN 154

QY 268 CLAPMAKALHDFAFGQKGLMTTIHAYTGDQMLDGPGRGDLRRRAGAANIVPNSTGAA 327

DB 155 CLAPMAVVLQKEFGVVEGLMTTIHAYTGDQNTLDGPHRKGLRRRARAALNIVPNSTGAA 214

QY 328 KAIGLVIPELNGKLDGAAQORVPVPTGSGVTVELVTLDKNVSVDEINAAKAAASNDSEGYTE 387

DB 215 KAIGLVIPELNGKLDGSAQRVPVATGSLTVELSVLVRPVTKEEINAAKAAASBSYGYNE 274

Search completed: September 15, 2004, 10:27:49
Job time : 43 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:11:48 ; Search time 25 Seconds
(without alignments)
933.097 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278

Sequence: 1 MKKITGILLALLAVIILSAC.....EMSYTAQLVRTLEYFAKIAK 448

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1650.5	72.5	335	1	G3P2_STRPY
2	1646.5	72.3	335	1	G3P2_STRPY3
3	1553.5	68.2	335	1	G3P2_STRPY
4	1280	56.2	337	1	G3P2_LACIA
5	1234	54.2	334	1	G3P2_CLOPA
6	1203	52.8	334	1	G3P2_CLOAB
7	1138.5	50.0	336	1	G3P2_STAEP
8	1132.5	49.7	336	1	G3P2_STAAM
9	955.5	41.9	338	1	G3P2_LACDE
10	941.5	41.3	333	1	G3P2_ECO57
11	940.5	41.3	333	1	G3P2_ECOLI
12	870.5	38.2	337	1	G3P2_MYCPN
13	854.5	37.5	337	1	G3P2_MYCGE
14	850	37.3	334	1	G3P2_MYCGL
15	844	37.1	334	1	G3P2_BACSU
16	840	36.9	334	1	G3P2_BACME
17	825	36.2	334	1	G3P2_BACST
18	820.5	36.0	335	1	G3P2_BORBU
19	818.5	35.9	332	1	G3P2_THEMA
20	793.5	34.8	342	1	G3P2_AQUAE
21	790	34.7	336	1	G3P2_STRCO
22	786.5	34.5	339	1	G3P2_MYCTO
23	783.5	34.4	339	1	G3P2_MYCLE
24	777.5	34.1	339	1	G3P2_MYCAV
25	763.5	33.5	330	1	G3P2_SALTY
26	759.5	33.3	330	1	G3P2_ECOLI
27	754	33.1	339	1	G3P2_HAEIN
28	743.5	32.6	330	1	G3P2_TRYBB
29	740	32.5	332	1	G3P2_RALSO
30	739	32.4	332	1	G3P2_STRAU
31	737	32.4	337	1	G3P2_RHURA
32	737	32.4	337	1	G3P2_MONAN
33	734	32.2	336	1	G3P2_SCHPO

34 732.5 32.2 337 1 G3P2_ANASP
35 729 32.0 337 1 G3P2_CLAPU
36 725 31.8 331 1 G3P2_THEAQ
37 724.5 31.8 336 1 G3P2_SYNY3
38 723 31.7 338 1 G3P2_NEUCR
39 722 31.7 337 1 G3P2_COCHU
40 721 31.7 337 1 G3P2_CURLU
41 719.5 31.6 333 1 G3P2_STRAE
42 717.5 31.5 330 1 G3P2_LEIME
43 717 31.5 335 1 G3P2_SCHPO
44 715 31.4 336 1 G3P2_ASPNG
45 713 31.3 336 1 G3P2_EMENI

P58554 anabaena sp
Q00584 claviiceps p
P00361 thermus aqu
P80505 synechocyst
P54118 neurospora
P29497 cochllobolu
P28844 curvularia
P54226 streptomyc
Q01558 leishmania
O43026 schizosacch
Q12552 aspergillus
P20445 emericella

ALIGNMENTS

RESULT 1

G3P2_STRPY STANDARD; PRT; 335 AA.
AC P50467;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)
DE (Plasminogen-binding protein) (Plasmin receptor).
GN GAP OR PLR OR GAPA OR SPY0274 OR SPYM18_0261.
OS Streptococcus pyogenes, and
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 186103;
RN [1]_TaxID=1314, 186103;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-73; 160-173 AND 186-216.
RC STRAIN=64/14;
RX MEDLINE=92355491; PubMed=1322883;
RA Lottenberg R., Broder C.C., Boyle M.D., Kain S.J., Schroeder B.L.,
RA Curtiss R. III;
RT "Cloning, sequence analysis, and expression in Escherichia coli of a
streptococcal plasmin receptor";
RL J. Bacteriol. 174:5204-5210(1992).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=SF370 / ATCC 700294 / Serotype M1;
RC MEDLINE=21192684; PubMed=11296296;
RX Ferretti J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus
pyogenes."; Acad. Sci. U.S.A. 98:4658-4663(2001).
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=MGAS8232 / Serotype M18;
RC MEDLINE=2197593; PubMed=11917108;
RX Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.D., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
group A Streptococcus strains associated with acute rheumatic fever
outbreaks"; Acad. Sci. U.S.A. 99:4668-4673(2002).
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
[4]
RN SEQUENCE OF 1-11; 20-30; 103-128; 162-171 AND 199-215.
RP STRAIN=JRS4 / Serotype M6;
RC Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,
RA VanBogelen R.A.;
RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes
proteins."; Submitted (MAY-2000) to Swiss-Prot.
RL Submitted (MAY-2000) to Swiss-Prot.
CC -I- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.

CC CC CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M95569; AAA26953.1; -;
CC EMBL; AE006494; AAK33346.1; -;
CC EMBL; AE009973; AAL97041.1; -;
CC HSP; P00362; IGD1.
CC InterPro; IPR000173; GAP_dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF02800; gpdh_C.1.
CC PRINTS; PR00078; G3PDHGRNASE.
CC PROSITE; PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD; Complete proteome.
CC INIT MET 0
CC BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
CC ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
CC (BY SIMILARITY).
CC CONFLICT 261 261 A -> S (IN REF. 1).
CC SEQUENCE 335 AA; 35811 MW; F06006EB253C8A3F CRC64;
Query Match 72.5%; Score 1650.5; DB 1; Length 335;
Best Local Similarity 79.0%; Pred. No. 1.2e-93;
Matches 332; Conservative 1; Mismatches 1; Indels 85; Gaps 1;
SQ
QY 29 VVKVINGFGRIQAFRRIONVEGVETRINDLTPDNLMLAHLLKYDTTQGRFDGTVGVK 88
Db 1 VVKVINGFGRIQAFRRIONVEGVETRINDLTPDNLMLAHLLKYDTTQGRFDGTVGVK 60
QY 89 EGGFVNGNFIKVSARDPENIDWATDGVVEIVLEALEGTVKDGDFVNGKFIKVSAREK 148
Db 61 EGGFVNGNFIKVSAREDE-----80
QY 149 DPEQIDWATDGVVEIVLEIDGTVEKGGFVNGQFVKVSAEREPANIDWATDGVVEIVLEA 208
Db 81 -----NIDWATDGVVEIVLEA 95
QY 209 TGFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVNTNHDLDGTETVIGSACTTNC 268
Db 96 TGFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVVNTNHDLDGTETVIGSACTTNC 155
QY 269 LAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGGLRRARAGAAINVPNSGAAK 328
Db 156 LAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGGLRRARAGAAINVPNSGAAK 215
QY 329 AGLVPELNGKLDGAAQRPVPTGTVELVTLTKNVSVDENIAMKAASNDSFGYTED 388
Db 216 AGLVPELNGKLDGAAQRPVPTGTVELVTLTKNVSVDENIAMKAASNDSFGYTED 275
QY 389 PIVSSDIVGVSGLSFLDQTQKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKTA 448
Db 276 PIVSSDIVGVSGLSFLDQTQKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKTA 335

RESULT 2
G3F_STRP3
ID - G3P_STRP3
AC Q8K8W9;

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)
DE (Plasminogen-binding protein) (Plasmin receptor).
GN GAP OR PLR OR SPVM3_0201 OR SPS0207.
OS Streptococcus pyogenes (serotype M3).
CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=138466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
EX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RA "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
EX MEDLINE=22683278; PubMed=12799345;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayaishi H., Hattori M., Hamada S.;
RA "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RL Genome Res. 13:1042-1055 (2003).
CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN (By similarity).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014140; AAM78808.1; -;
CC EMBL; AP005141; BAC63302.1; -;
CC InterPro; IPR000173; GAP_dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh_C.1.
CC Pfam; PF02800; gpdh_C.1.
CC PRINTS; PR00078; G3PDHGRNASE.
CC TIGRFS; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD; Complete proteome.
CC INIT MET 0
CC BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
CC ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
CC (BY SIMILARITY).
CC SEQUENCE 335 AA; 35841 MW; 4DCB76F382F6F98 CRC64;
Query Match 72.3%; Score 1646.5; DB 1; Length 335;
Best Local Similarity 79.0%; Pred. No. 2.1e-93;
Matches 332; Conservative 1; Mismatches 2; Indels 85; Gaps 1;
SQ
QY 29 VVKVINGFGRIQAFRRIONVEGVETRINDLTPDNLMLAHLLKYDTTQGRFDGTVGVK 88

Db 1 VVKVINGFORIGRLAFRRIONIEGVETVTRINDLTDNMLAHLKLYDTTQGRFDGTVEVK 60
QY 89 EGGFEVNGNFIKVSAERDPENIDWATDGVBRIVLEALEGTVEVKGDPVNGKFKIVSAEK 148
Db 61 EGGFEVNGNFIKVSAERDPE-----80
QY 149 DPEQIDWATDGVBRIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVBRIVLEA 208
Db 81 -----NIDWATDGVBRIVLEA 95
QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGNDVKTWVFNTHDILDTGTTVISGASCTTNC 268
Db 96 TGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTWVFNTHDILDTGTTVISGASCTTNC 155
QY 269 LAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGGLRRARAGAAANIVPNSGAAK 328
Db 156 LAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGGLRRARAGAAANIVPNSGAAK 215
QY 329 AIGLVIPELNGKLDGAAQRPVPTGSGVTELVVTLDDKNVSDVDEINAAKAAASNDGFGYTED 388
Db 216 AIGLVIPELNGKLDGAAQRPVPTGSGVTELVVTLDDKNVSDVDEINAAKAAASNDGFGYTED 275
QY 389 PIVSSDIVGVSYSGLFDATQTKWMEVDGSQLVKKVSVWYDNMSYTAQLVRLTLEYFAKIAK 448
Db 276 PIVSSDIVGVSYSGLFDATQTKWMEVDGSQLVKKVSVWYDNMSYTAQLVRLTLEYFAKIAK 335
RESULT 3
G3P_STREQ
ID_G3P_STREQ STANDARD; PRT; 335 AA.
AC Q59506;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)
DE (Plasminogen-binding protein) (Plasmin receptor).
GN GAP OR GAPC.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H46A;
RX MEDLINE=96305364; PubMed=8706717;
RA Gase K., Gase A., Schirmer H., Malke H.;
RT "Cloning, sequencing and functional overexpression of the
RT Streptococcus equisimilis H46A gapC gene encoding a
RT glycerinaldehyde-3-phosphate dehydrogenase that also functions as a
RT plasmin(ogen)-binding protein. Purification and biochemical
RT characterization of the protein.";
RL Eur. J. Biochem. 239:42-51(1996).
CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glycerinaldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X97788; CAA66377.1; -;
CC DR EMBL; Y12602; CAA73174.1; -;
CC DR PIR; S71350; S71350.
CC HSSP; P00362; IGD1.

DR InterPro: IPR000173; GAP dhydrogenase.
DR InterPro: IPR006424; GAPDH-I.
DR Pfam: PF00044; gpdh; 1.
DR Pfam: PF02800; gpdh; C; 1.
DR PRINTS: PR00078; G3PDHGRNASE.
DR TIGRFAMS: TIGR01534; GAPDH-I; 1.
DR PROSITE: PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD.
FT INIT_MET 0 0
FT BINDING 151 151 BY SIMILARITY.
FT ACT_SITE 178 178 GLYCERALDEHYDE 3-PHOSPHATE (BY
FT SIMILARITY).
FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
SQ SEQUENCE 335 AA; 35739 MW; FE7ACFD7663E46 CRC64;
Query Match 58.2%; Score 1553.5; DB 1; Length 335;
Best Local Similarity 76.0%; Pred. No. 9.5e-88;
Matches 319; Conservative 2; Mismatches 14; Indels 85; Gaps 1;
QY 29 VVKVINGFORIGRLAFRRIONIEGVETVTRINDLTDNMLAHLKLYDTTQGRFDGTVEVK 88
Db 1 VVKVINGFORIGRLAFRRIONIEGVETVTRINDLTDNMLAHLKLYDTTQGRFDGTVEVK 60
QY 89 EGGFEVNGNFIKVSAERDPENIDWATDGVBRIVLEALEGTVEVKGDPVNGKFKIVSAEK 148
Db 61 EGGFEVNGNFIKVSAERDPE-----80
QY 149 DPEQIDWATDGVBRIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVBRIVLEA 208
Db 81 -----NIDWATDGVBRIVLEA 95
QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGNDVKTWVFNTHDILDTGTTVISGASCTTNC 268
Db 96 TGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTWVFNTHDILDTGTTVISGASCTTNC 155
QY 269 LAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGGLRRARAGAAANIVPNSGAAK 328
Db 156 LAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGGLRRARAGAAANIVPNSGAAK 215
QY 329 AIGLVIPELNGKLDGAAQRPVPTGSGVTELVVTLDDKNVSDVDEINAAKAAASNDGFGYTED 388
Db 216 AIGLVIPELNGKLDGAAQRPVPTGSGVTELVVTLDDKNVSDVDEINAAKAAASNDGFGYTED 275
QY 389 PIVSSDIVGVSYSGLFDATQTKWMEVDGSQLVKKVSVWYDNMSYTAQLVRLTLEYFAKIAK 448
Db 276 PIVSSDIVGVSYSGLFDATQTKWMEVDGSQLVKKVSVWYDNMSYTAQLVRLTLEYFAKIAK 335
RESULT 4
G3P_LACIA
ID_G3P_LACIA STANDARD; PRT; 337 AA.
AC P52987;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR LL0559.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LM0230.
RX MEDLINE=95291425; PubMed=7773380;
RA Cantilla M.R., Hillier A.J., Davidson B.E.;
RT "Lactococcus lactis glycerinaldehyde-3-phosphate dehydrogenase gene,
RT gap: further evidence for strongly biased codon usage in glycolytic
RT pathway genes";
RL Microbiology 141:1027-1036(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;

RA Bolotin A., Wincker P., Mauger S., Jallion O., Malarne K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium *Lactococcus*
 RL *lactis* ssp. *lactis* IL1403";
 Genome Res. 11:731-753(2001).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L36907; AAC1453.1; -;
 CC DR EMBL; AE006290; AA04857.1; -;
 CC DR PIR; G86694; G86694.
 CC DR HSP; P17721; IHG.
 CC DR InterPro; IPR000173; GAP_dhhydrogenase.
 CC DR InterPro; IPR006424; GAPDH-I.
 CC DR Pfam; PF00044; gpdh; 1.
 CC DR Pfam; PF02800; gpdh.C; 1.
 CC DR PRINTS; PR00078; G3PDHGRGNASE.
 CC DR TIGRfams; TIGR01534; GAPDH-I; 1.
 CC DR PROSITE; PS00071; GAPDH; FALSE NEG.
 CC KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
 CC FT GLYCERALDEHYDE 3-PHOSPHATE
 CC FT BINDING 152 152
 CC FT (BY SIMILARITY).
 CC FT ACT_SITE 179 179
 CC FT ACTIVATES THIOL GROUP DURING CATALYSIS
 CC FT (BY SIMILARITY).
 CC FT CONFLICT 143 143
 CC FT T -> S (IN REF. 1).
 CC FT
 CC SQ SEQUENCE 337 AA; 36057 MW; 17BB8C6AAEFF589D CRC64;
 CC
 CC Query Match 56.2%; Score 1280; DB 1; Length 337;
 CC Best Local Similarity 61.4%; Pred. NO. 4.2e-71;
 CC Matches 29; Conservative 31; Mismatches 46; Indels 86; Gaps 2;
 CC
 CC QY 28 MVKVGINGFGRIGRLAFRIQNVGEVETRLNDLPNNLAHLKYDTTQGRFGDTVEV 87
 CC 1 MVKVGINGFGRIGRLALRRIQVEGVEVAHNDLTPNLAHLKYDTTQGRFGDTVEV 60
 CC
 CC QY 88 KEGGFVNGNFVKSAERPDENIDWATDGVVEIVLEAGTVEVKGGFDVNGKFIKVSAB 147
 CC 61 KEGGFVNGKFKVTAERNPE
 CC
 CC QY 148 KDPEQIDWATDGVVEIVLEIDGTVVEKGGFVNGKFIKVSAB 147
 CC 61 KEGGFVNGKFKVTAERNPE
 CC
 CC QY 148 KDPEQIDWATDGVVEIVLEIDGTVVEKGGFVNGKFIKVSAB 147
 CC 61 KEGGFVNGKFKVTAERNPE
 CC
 CC QY 82
 CC
 CC QY 208 ATSPFAKKEAPKHLHANGAKKVVITAPGNDVKTVVNTNHDLDGTETVLSGASCTTN 267
 CC 96 ATGFTATKEAKSKLHPGAKKVLITAPGNDVKTVVNTNHDLDGTETVLSGASCTTN 155
 CC
 CC QY 268 CLAPMAKALHDAFGTQKGLMTTHAYTGDMQLDGHRRGDLRRRAGAANTVPNSTGAA 327
 CC 156 SLAPMADALKNFGVKGGTMTTTHSPTVTDQMTLDGPHRGDFRRARAANIVPASSGAA 215
 CC
 CC QY 328 KAIGVIVPBLNGKLDGAQRVPVPTGVTTELVDKNSVDDEINAAKKAASNDSPGYTE 387
 CC 216 KAIGLVLPGLSKMGHAQRVSTPGTITELVTLEKVVDEINAAKKAASNDSPGYTNV 275
 CC
 CC QY 368 DVISSDIVGVSGISFDATQTKMVEV-DGSLVKVSVWYDNMSYTAQLVETLEYFAKI 446
 CC 276 DEIVSSDIIGMAYGISFDATLTETVTDLKGGLVKTAAWYDNMSFTAQLTLEYFAKI 335
 CC
 CC QY 447 AK 448

Db 336 AK 337
 ||
 RESULT 5
 G3P_CLOPA STANDARD; PRT; 334 AA.
 AC Q99309;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) (CP
 DE 17/CP 18).
 GN GAP.
 OS Clostridium pasteurianum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OC NCBI_TaxID=1501;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oster T., Assobhei O., Scherrer S., Branlant G., Branlant C.;
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP PARTIAL SEQUENCE OF 1-26.
 RC STRAIN=W5;
 RX MEDLINE=98291870; PubMed=9629918;
 RA Flengsrud R., Skjeldal L.;
 RT "Two-dimensional gel electrophoresis separation and N-terminal
 RL sequence analysis of proteins from *Clostridium pasteurianum* W5.";
 RL Electrophoresis 19:802-806(1998).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X72219; CAA51020.1; -;
 CC DR PIR; S34254; S34254.
 CC DR HSP; P00362; 1GDI.
 CC DR InterPro; IPR000173; GAP_dhhydrogenase.
 CC DR InterPro; IPR006424; GAPDH-I.
 CC DR Pfam; PF00044; gpdh; 1.
 CC DR Pfam; PF02800; gpdh.C; 1.
 CC DR PRINTS; PR00078; G3PDHGRGNASE.
 CC DR TIGRfams; TIGR01534; GAPDH-I; 1.
 CC DR PROSITE; PS00071; GAPDH; 1.
 CC KW Glycolysis; Oxidoreductase; NAD.
 CC FT BINDING 150 150
 CC FT GLYCERALDEHYDE 3-PHOSPHATE.
 CC FT ACTIVATES THIOL GROUP DURING CATALYSIS.
 CC FT ACT_SITE 177 177
 CC SQ SEQUENCE 334 AA; 36078 MW; D15905D0DA7F62E7 CRC64;
 CC
 CC Query Match 54.2%; Score 1234; DB 1; Length 334;
 CC Best Local Similarity 59.3%; Pred. No. 2.6e-68;
 CC Matches 249; Conservative 34; Mismatches 51; Indels 86; Gaps 3;
 CC
 CC QY 29 VVKVGINGFGRIGRLAFRIQNVGEVETRLNDLPNNLAHLKYDTTQGRFGDTVEV 88
 CC 1 MTKVAINGFGRIGRLALRRILEVPGVEVAINDLTPNLAHLKYDTTQGRFGDTVEV 60
 CC
 CC QY 89 EGGFVNGNFVKSAERPDENIDWATDGVVEIVLEAGTVEVKGGFDVNGKFIKVSAB 148
 CC 61 EGGFVNGKFKVTAERNPE
 CC
 CC Db 61 EGGFVNGKFKVTAERNPE


```

QY 149 DPEQIDWATGCVGVEIVLEIDGTVEVKEGGFVNGQFVKVSAEREPAIDWATGVEIVLEA 208
Db 85 -----GEL-----GIDVLEEC 95
QY 209 TSFAKKEAEKHLHANGAKKVVITAPGNDVKTVWFNTNHDILDTGTVISGASCTTNC 268
Db 96 TGFTTKEKAEAHVRA-GAKKVVISAPAGNDLKTIVFVNNEDLDGTETVISGASCTTNC 154
QY 269 LAPMAKALHDAPFGIQKGLMTTHAYTGDQMLDGHRRGDLRRRAGAANTVPSNSTGA 328
Db 155 LAPMAKVLNDKFGIEKGPMTHAYTNDQNTLDGPHRKGDFERRARAAVSIIPNSTGA 214
QY 329 AIGLVIPELNGKLDGAAQRPVPGSVTVLVTLDKNVSVDEINNAKKAASNDSEGYTED 388
Db 215 AIAQVIPLEKGLDGNAGQRPVPTGTVTELTVLKNVTVBEINNAKKEANESFGYTED 274
QY 389 PIVSSDIYGVSGSLFDATQTKMVEVDSQVVKVSVWYDNEVSYTAQLVRLTLEYFAKIAK 448
Db 275 EIVSADVVGISYGLSFDATLTKIVDVDSQVVKVSVWYDNEVSYTSLVRLTLEYFAKIAK 334

RESULT 6
G3P_CLOAB
ID G3P_CLOAB STANDARD; PRT; 334 AA.
AC 052631.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) .
GN GAP OR GACP OR CAC0709.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=99192446; PubMed=10463150;
RA Schreiber W., Durre P.;
RT "The glyceraldehyde-3-phosphate dehydrogenase of Clostridium
RT acetobutylicum: Isolation and purification of the enzyme, and
RT sequencing and localization of the gap gene within a cluster of other
RT glycolytic genes.";
RL Microbiology 145:1839-1847(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF043386; AAC13160.1; -
EMBL; AE007586; AAK78686.1; -

```

```

PIR: C96987; C96987.
HSP; P17721; IHG.
InterPro; IPR000173; GAP_dhhydrogenase.
InterPro; IPR006424; GAPDH-1.
Pfam; PF00044; gpdh; 1.
Pfam; PF02800; gpdh_C; 1.
PRINTS; PR00078; G3PDHGRGNASE.
TIGRFAMs; TIGR01534; GAPDH-I; 1.
PROSITE; PS00071; GAPDH; 1.
Glycolysis; Oxidoreductase; NAD; Complete proteome.
KW GLYCERALDEHYDE 3-PHOSPHATE.
FT BINDING 150 150
FT ACT SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 334 AA; 35850 MW; 10C52A174BE789B5 CRC64;

Query Match 52.8%; Score 1203; DB 1; Length 334;
Best Local Similarity 57.4%; Pred. No. 2e-66;
Matches 241; Conservative 41; Mismatches 52; Indels 86; Gaps 3;

QY 29 VVKVINGFGGRIGLAFRRIQNVGVETRLNDLTPNMLAHLKYDTTQGRFGDTVEVK 88
Db 1 MAKTAINGFGGRIGLALRRILEVEGLEVVAINDLTDKMLAHLKYDSSQGRFNGEIVK 60
QY 89 EGGFEVNGNFIVSAERDPENIDWATGVEIVLEALECTVEVKGDFVNGKFIKVSAAK 148
Db 61 EGAFVNGKEVKVFAEADPEKLPWG-----DL----- 87
QY 149 DPEQIDWATGVEIVLEIDGTVEVKEGGFVNGQFVKVSAEREPAIDWATGVEIVLEA 208
Db 88 -----GIDVLEEC 95
QY 209 TSFAKKEAEKHLHANGAKKVVITAPGNDVKTVWFNTNHDILDTGTVISGASCTTNC 268
Db 96 TGFTTKEKAEAHVRA-GAKKVVISAPAGNDLKTIVFVNNEDLDGTETVISGASCTTNC 154
QY 269 LAPMAKALHDAPFGIQKGLMTTHAYTGDQMLDGHRRGDLRRRAGAANTVPSNSTGA 328
Db 155 LAPMAKVLNDKFGIEKGPMTHAYTNDQNTLDGPHRKGDLRRARAAVSIIPNSTGA 214
QY 329 AIGLVIPELNGKLDGAAQRPVPGSVTVLVTLDKNVSVDEINNAKKAASNDSEGYTED 388
Db 215 AIAQVIPLEKGLDGNAGQRPVPTGTVTELTVLKNVTVBEINNAKKEANESFGYTED 274
QY 389 PIVSSDIYGVSGSLFDATQTKMVEVDSQVVKVSVWYDNEVSYTAQLVRLTLEYFAKIAK 448
Db 275 EIVSADVVGISYGLSFDATLTKIVDVDSQVVKVSVWYDNEVSYTSLVRLTLEYFAKIAK 334

RESULT 7
G3P1_STAEP
ID G3P1_STAEP STANDARD; PRT; 336 AA.
AC Q8CPY5;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH 1) .
GN GAPA OR GAP OR S05057.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
RT Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593(2003).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).

```

RESULT 8

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AJ000339; CAA04014.1; -;
 CC PIR; T09633; T09633.
 CC HSP; P17721; 1HDG.
 CC InterPro; IPR000173; GAP_dhhydrogenase.
 CC InterPro; IPR008424; GAPDH-I.
 CC Pfam; PF00044; gpdh; 1.
 CC PRINTS; PR00078; G3PDHGNASE.
 CC TIGRfams; TIGR01534; GAPDH-I; 1.
 CC PROSITE; PS00071; GAPDH; FALSE_NEG.
 KW Glycolysis; NAD; Oxidoreductase.
 FT BINDING 156 156 GLYCERALDEHYDE 3-PHOSPHATE
 FT ACT_SITE 183 183 (BY SIMILARITY).
 FT ACTIVATES THIOL GROUP DURING CATALYSIS
 FT (BY SIMILARITY).
 SQ SEQUENCE 338 AA; 36564 MW; DF97B8C9CA4F7DFA CRC64;
 Query Match 41.9%; Score 955.5; DB 1; Length 338;
 Best Local Similarity 47.3%; Pred. No. 2.3e-51;
 Matches 200; Conservative 46; Mismatches 88; Indels 89; Gaps 4;

QY 28 MVKVGINGFGRIQLAFRR---QNVGVEVTRINDLTPNMLAHLKYDTTQGRFDG 83
 DB 1 MTVKIGINGFGRIQLAFRRIMDLGEETKQIEVVAINDLTPMLAHLKYDSTHGTFDH 60
 QY 84 TVEVKEGFGFVNGNFIKVSARDPENTIDWATDGVIEVLEALEGTVKDGDFVNGKFIK 143
 DB 61 EVSATEDSLVDGKKRYVAEPQAQNPW-----VKN----- 92
 QY 144 VSAEKPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGV 203
 DB 93 -----DGV 96
 QY 204 IVLEATSEFAKKEAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTETVISGAS 263
 DB 97 FVLECTGFTYTSKAKSAHLDA-GAKRVLISAPAGNDLKTIVISVNDLTADDTTIVSAGS 155
 QY 264 CTNCLAPMAKALHDAFGIQKGLMTTHAYTGDMILDPHRRGDLRRARAGAAIVPNS 323
 DB 156 CTNLSLAPMANALNKEFGIQVGTWTHAYTATQKVLDPGRGNFRNARAENIIPHS 215
 QY 324 TGAAKAIGLVIPDLNGKLDGAAQRPVPTGVTSLVTLTKNSVDEINAAKKAASNDSF 383
 DB 216 TGAAKAIGLVIPDLNGKLDGHAQRPVVKDGSSETLVTLTKKVTAEVNAAMKYESPSF 275
 QY 384 GYTEDPIVSSDIVGVSGSLPDATQTKVMEVDGSLQVSVSWYDNEMSVTAOLVTLLEYF 443
 DB 276 AYNADQIVSTDVLGWTAGSIFDPTQTQITAGDKQLVKTVAWYDNEYSEFTQMVATLLHF 335
 QY 444 AKI 446
 DB 336 ATL 338

RESULT 10
 ID G3P3 ECO57 STANDARD; PRT; 333 AA.
 AC P58072; Rel. 40, Created
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase C (EC 1.2.1.12) (GAPDH-C).
 GN GAPC OR Z2304 OR ECS2022.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.

CC EMBL; AJ000339; CAA04014.1; -;
 CC PIR; T09633; T09633.
 CC HSP; P17721; 1HDG.
 CC InterPro; IPR000173; GAP_dhhydrogenase.
 CC InterPro; IPR008424; GAPDH-I.
 CC Pfam; PF00044; gpdh; 1.
 CC PRINTS; PR00078; G3PDHGNASE.
 CC TIGRfams; TIGR01534; GAPDH-I; 1.
 CC PROSITE; PS00071; GAPDH; FALSE_NEG.
 KW Glycolysis; NAD; Oxidoreductase.
 FT BINDING 156 156 GLYCERALDEHYDE 3-PHOSPHATE
 FT ACT_SITE 183 183 (BY SIMILARITY).
 FT ACTIVATES THIOL GROUP DURING CATALYSIS
 FT (BY SIMILARITY).
 SQ SEQUENCE 338 AA; 36564 MW; DF97B8C9CA4F7DFA CRC64;
 Query Match 41.9%; Score 955.5; DB 1; Length 338;
 Best Local Similarity 47.3%; Pred. No. 2.3e-51;
 Matches 200; Conservative 46; Mismatches 88; Indels 89; Gaps 4;

QY 28 MVKVGINGFGRIQLAFRR---QNVGVEVTRINDLTPNMLAHLKYDTTQGRFDG 87
 DB 1 MAVKVAINGFGRIQLAFRR---QNVGVEVTRINDLTPNMLAHLKYDTTQGRFTGEVEV 60
 QY 88 KEGGFEVNGNFIKVSARDPENTIDWATDGVIEVLEALEGTVKDGDFVNGKFIKVSAAE 147
 DB 61 VDGGFRVNGKEVSEFSDASKLPK-----DLN----- 89
 QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 DB 90 -----IDV 95
 QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTETVISGASCTTN 267
 DB 96 CTGFYTDKKAQAHLA-GAKKRVLISAPATGDLKTIVNTNHDILDTGTETVISGASCTTN 154
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDPHRRGDLRRARAGAAIVPNSGAA 327
 DB 155 SLAPVAKVINDFGVLEGLMTTHAYTGDMILDPHRRGDLRRARAGAAIVPNSGAA 214
 QY 328 KAIGLVIPDLNGKLDGAAQRPVPTGVTSLVTLTKNSVDEINAAKKAASNDSPGYT 386
 DB 215 KAIGKVIPEIDGKLDGGAQRPVATGSLTSLVTLTKKVTAEVNAAMKKAASNESFGYT 274
 QY 387 EDPVSSDIVGVSGSLPDATQTKVMEVDGSLQVSVSWYDNEMSVTAOLVTLLEYFAKI 446
 DB 275 EDEIVSSDVGVSGSLPDATQTRVMSVGDRLQVKAAYDNEMSVTAOLVTLLEYFAKI 334
 QY 447 AKI 448
 DB 335 SK 336

RESULT 9
 ID G3P3 LACDE STANDARD; PRT; 338 AA.
 AC Q32755;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAP.
 OS Lactobacillus delbrueckii (subsp. bulgaricus).
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 OX NCBI_TaxID=1585;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B107;
 RX MEDLINE=98240227; PubMed=9579064;
 RA Branny P., Delatorre F., Garell J.R.;
 RT "An operon encoding three glycolytic enzymes in Lactobacillus
 RT delbrueckii subsp. bulgaricus: glyceraldehyde-3-phosphate
 RT dehydrogenase, phosphoglycerate kinase and triosephosphate
 RT isomerase".
 RL Microbiology 144:905-914 (1998).
 CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -!- PATHWAY: Second phase of glycolysis; first step.
 CC -!- SUBUNIT: Homotetramer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta B.T., Potamouis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe I.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AA05364; AAG56359.1; -;
 DR EMBL; AF002557; BAB35445.1; -;
 DR PIR; C85737; C85737.
 DR PIR; F90881; F90881.
 DR HSSP; P17721; 1HDG.
 DR InterPro; IPR000173; GAP_dhdrogenase.
 DR InterPro; IPR006424; GAPDH-I.
 DR Pfam; PF00044; gpdc; 1.
 DR Pfam; PF02800; gpdc; 1.
 DR PRINTS; PR00078; G3PDHRCNASE.
 DR TIGRfam; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
 FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE
 FT (BY SIMILARITY).
 FT ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS
 FT (BY SIMILARITY).
 FT FT
 SQ SEQUENCE 333 AA; 35763 MW; A2F77CB2E773B64C CRC64;
 Query Match 41.3%; Score 941.5; DB 1; Length 333;
 Best Local Similarity 47.5%; Pred. No. 1.6e-50;
 Matches 199; Conservative 53; Mismatches 76; Indels 91; Gaps 7;
 QY 31 KVGINGFGIGRLAFRRIONVE-GVEVTRINDLTPNMAHLKYDTTGGDFGVVEVKE 89
 Db 3 KVGINGFGIGRLVRLLEVNIDVVAINDLTPKILAYLLKHSNYPGPF----- 54
 QY 90 GGFVNGNFIKVSARDPENIDWATDGVLEIVLEALEGTVVEVKGDFVNGKFIKVSARD 149
 Db 55 ----- 54
 QY 150 PQIDWATDGVETVLBDTGVKEGFEVNGQFVKVSAEREPANIDWATDGVETVLEAT 209
 Db 55 ----PWSVDFTEDSLIVD-----GKSIIVAEKAKNIPWAKAGAEIIVECT 97
 QY 210 SFFAKKEAAEKHLHANGAKKVVITAPGNGNDKVVTVFNTNHDILDTGTVTISGASCTTNCL 269

Db 98 GFTTSAEKSOAHLDA-GAKVLISAAG-EMKTIIVKVNDDTLDGNDTIVSVASCTTNCL 155
 QY 270 APMAKALHDAFGIQKGLMTTHAYTDQDMILDPGHRGGLRRARAGAANVPNSTGAAGA 329
 Db 156 APMAKALHDSFGIEVGTMTTHAYTGTQSLVDGP-RGKDLRASRAAAENIIPHTTGAAGA 214
 QY 330 IGAIVPELNGKLDGAORVPVPTGSGVTTELWVLDKNVSVDEINAAKAA--SNDSEGYTE 387
 Db 215 IGLVPELNGKLDGAORVPVPTGSGVTTELWVLDKNVSVDEINAAKAA--SNDSEGYTE 274
 QY 388 DPTVSSDIVGSYSLFDATQTKVMEVDGSQLVKVSYVDNEMSYTAQIVRTLEYPAKI 446
 Db 275 EIRVSSDIIGSHFGSVFDATQTEITAVGLQLVKTVAWVDNEYGFVTLRTLEKPAKL 333
 RESULT 11
 G3P3 ECOLI STANDARD; PRT: 333 AA.
 ID G3P3 ECOLI STANDARD; PRT: 333 AA.
 AC P33858; P76094; P78062; P78291; Q03850; Q63208;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase C (BC 1.2.1.12) (GAPDH-C).
 GN GAPC OR B1416/B1417.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RX SEQUENCE FROM N.A.
 RC STRAIN=River isolate, and Clinical isolate;
 RX MEDLINE=98283700; PubMed=9622357;
 RA Espinosa-Urgel M., Kolter R.;
 RT "Escherichia coli genes expressed preferentially in an aquatic
 RT environment";
 RL Mol. Microbiol. 28:325-332(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Hidalgo E., Limon A., Aguilar J.;
 RT "A second Escherichia coli gene with similarity to gapA";
 RL Microbiologia 12:99-106(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=972511357; PubMed=9097039;
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 RA Kasa H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sampei G., Seki Y., Sivasubdam S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map";
 RL DNA Res. 3:363-377(1996).
 RN [5]
 RP SEQUENCE OF 306-332 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=92011371; PubMed=1917845;
 RA Hidalgo E., Chen Y.-M., Lin E.C.C., Aguilar J.;
 RT "Molecular cloning and DNA sequencing of the Escherichia coli K-12
 RT ald gene encoding aldehyde dehydrogenase";

J. Bacteriol. 173:6118-6123(1991).
 [6] SEQUENCE OF 1-13 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=88232416; PubMed=2836696;
 RA Nakamura H., Murakami H., Yamato I., Anraku Y.;
 RT "Nucleotide sequence of the cybB gene encoding cytochrome b561 in
 Escherichia coli K12.";
 RL Mol. Gen. Genet. 212:1-5(1988).
 [7] SEQUENCE OF 1-165 FROM N.A.
 RP Krawetz S.A.;
 RA Submitted (OCT-1990) to the EMBL/GenBank/DBSJ databases.
 RL [8]
 RN SEQUENCE OF 1-52 FROM N.A.
 RP MEDLINE=87053179; PubMed=3780374;
 RX Krawetz S.A., Connor W., Cannon P.D., Dixon G.H.;
 RA "A vector-primer-cloner-sequencer plasmid for the construction of cDNA
 libraries: evidence for a rat glyceraldehyde-3-phosphate
 dehydrogenase-like mRNA and a ferritin mRNA within testis.";
 RT DNA 5:427-435(1986).
 RL [9]
 CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate +
 NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -!- PATHWAY: Second phase of glycolysis; first step.
 CC -!- SUBUNIT: Homotrimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 dehydrogenase family.
 CC -!- CAUTION: IN THE K12 STRAIN THIS GENE IS DISRUPTED BY A STOP CODON
 AND A FRAMESHIFT. IT SEEMS TO BE INTACT IN A NUMBER OF WILD
 STRAINS.
 CC -!- CAUTION: Ref.7 and Ref.8 sequences were originally thought to
 originate from rat.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L09067; AAA23856.1; ALT FRAME.
 DR EMBL; AE000239; AAC74498.1; ALT_FRAME.
 DR EMBL; AE000239; AAC74499.1; ALT_FRAME.
 DR EMBL; D90780; BAA15033.1; ALT_FRAME.
 DR EMBL; D90781; BAA15038.1; ALT_FRAME.
 DR EMBL; M64541; -; NOT_ANNOTATED_CDS.
 DR EMBL; M64541; -; NOT_ANNOTATED_CDS.
 DR EMBL; X54798; CAA38569.1; -.
 DR EMBL; M14166; AAA41178.1; -.
 DR HSSP; F17721; 1HDG.
 DR EcoGene; EG12103; gapC.
 DR InterPro; IPR000173; GAP_dhhydrogenase.
 DR InterPro; IPR006424; GAPDH-I.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh_C; 1.
 DR PRINTS; PR00078; G3PDHGRNASE.
 DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
 FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE
 (BY SIMILARITY).
 FT ACT_SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS
 (BY SIMILARITY).
 FT CONFLICT 39 39 K -> Y (IN REF. 7 AND 8).
 FT SEQUENCE 333 AA; 35649 MW; E88223297376B0A0 CRC64;
 Query Match 41.3%; Score 940.5; DB 1; Length 333;
 Best Local Similarity 47.5%; Pred. No. 1.9e-50;
 Matches 199; Conservative 53; Mismatches 76; Indels 91; Gaps 7;
 31 KVGINGFGRIGRLVGLLEVKSNIDVVAINDLSPKILAYLLKHSNYPF----- 54

Db 3 KVGINGFGRIGRLVGLLEVKSNIDVVAINDLSPKILAYLLKHSNYPF----- 54
 Qy 90 GGEVNGNFIVKSAERDPENIDWATDGVVEIVLEALEGTVEVKDGGDFVNGKFIKVSAEKD 149
 Db 55 ----- 54
 Qy 150 PEQIDWATDGVVEIVLEIDGTVEVEKGGFVNGQFVKVSAEREPANIDWATDGVVEIVLEAT 209
 Db 55 -----PWSVDFTDSLIVD-----GKSIAYVAEKEAKNIPKAKGAEIIVECT 97
 Qy 210 SFFAKAEAEKHLHANGAKKVVITAPGNDVKTVFVFNTHDILDTETVIGASCTTNCL 269
 Db 98 GFYISAEKSOAHLDA-GAKKVLISAPAG-EMKTIYVNVNDTLDGNDTIVSVASCTTNCL 155
 Qy 270 APMALHDAFGIQKGLMTTTHAYTGQMDLGDGPHRGDLRRARAGAAIIVPNSTGAAGA 329
 Db 156 APMALHDSFGIEVGVTTHAVTGTQSLVDGP-RGKDLRASRAAAENIIPHTTGAAGA 214
 Qy 330 IGLVPELNGKLDGAAGRPVPTGVTGLVTLVTLKNSVDSDEINAAKAA--SNDSFGYTE 387
 Db 215 IGLVPELNGKLDGAAGRPVPTGVTGLVTLVTLKNSVDSDEINAAKAA--SNDSFGYTE 274
 Qy 388 DVISSDIIVGVSGLFDTOTKMEVDGSQLVKVSWYDNMSYTAOLVTLTLEYFAKI 446
 Db 275 EEIVSSDIIGSHFGSVFDTQTEITAVGDLQVKTVAWYDNEYGFVTLITLKEFAKL 333
 RESULT 12
 G3P_MYCPN STANDARD; PRT; 337 AA.
 ID G3P MYCPN STANDARD; PRT; 337 AA.
 AC P75358;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAPA OR GAP OR MP430 OR MP411.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelreich R., Hilbert H., Piagens H., Pirkl E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 pneumoniae.";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -!- PATHWAY: Second phase of glycolysis; first step.
 CC -!- SUBUNIT: Homotrimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 dehydrogenase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE000040; AAB96059.1; -.
 DR PIR; S73737; S73737.
 DR HSSP; F17721; 1HDG.
 DR InterPro; IPR000173; GAP_dhhydrogenase.
 DR InterPro; IPR006424; GAPDH-I.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh_C; 1.
 DR PRINTS; PR00078; G3PDHGRNASE.


```

AC Q01651;
DT 01-APR-1993 (Rel. 25, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR CGL1588.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13059 / AS019;
RX MEDLINE=93015645; PubMed=1400158;
RA Eikmanns B.J.;
RT "Identification, sequence analysis, and expression of a
RT Corynebacterium glutamicum gene cluster encoding the three glycolytic
RT enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate
RT kinase, and triosephosphate isomerase."
RN J. Bacteriol. 174:6076-6086(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X59403; CAA42045.1; -.
CC EMBL; AP005279; BAB98981.1; -.
CC F1; A43260; A43260.
CC HSSP; P00362; 1GDI.
CC InterPro; IPR000173; GAP_dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh, 1.
CC Pfam; PF02800; gpdh_C, 1.
CC PRINTS; PR00078; G3PDHGRGNASE.
CC TIGRFAMs; TIGR01534; GAPDH-I, 1.
CC PROSITE; PS00071; GAPDH, 1.
CC GlycoSys; Oxidoreductase; NAD; Complete proteome.
KW BINDING 153 153 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT_SITE 180 180 ACTIVATES THIOL GROUP DURING CATALYSIS.
FT CONFLICT 25 26 SD -> NG (IN REF. 1).
FT CONFLICT 333 334 KL -> QALN (IN REF. 1).
SQ SEQUENCE 334 AA; 36045 MW; 33792AF65FA90FA7 CRC64;

Query Match 37.3%; Score 850; DB 1; Length 334;
Best Local Similarity 44.7%; Pred. No. 6e-45;
Matches 189; Conservative 41; Mismatches 95; Indels 98; Gaps 7;

QY 28 MVVKGINGFGRIGLAFPRRI-QNVEGVETVTRINDLPNMLAHLIKYDTTQGRFDGIVE 86
Db 1 MTRVINGFGRIGRNFFRAVLESDLEWAVNLDLTKNTLTLKSPDSIMGRLGQVE 60
QY 87 VKEGFEVNGNFIKVAERDPENIDWATDGVBEIVLEALEGTVEVKGDFVNGKFIKVA 146
Db 61 YDDSDSITVGKRIAVVAERDPK----- 82
QY 147 EKDPQIDWATDGVBEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVBEIVL 206

```

```

Db 83 -----NLDWAHNVDIVI 95
QY 207 EATSFAXKEAAEKHLHANGAKKVVITAPGNDVKVTVFNTNHDLD-GTETVISGASCT 265
Db 96 ESTGFTDANAAKAHTEA-GAKKVIISAPASNEDATFVYGVNHSYDPENHNVISGASCT 154
QY 266 TNCCLAPMAKALHDAPGQKGLMTTTHAYTGDMLDGPHPGGDLRRARAGAANIVPNSGT 325
Db 155 TNCCLAPMAKVLNDKFGLENGMTTTHAYTGDQRLHDAPHR--DLRRARAANVIVPTSTG 212
QY 326 AAKAIGLIVIPELNGKLDGAARQVPEPTGVTSLVTLVTLTKNVSVDEINAAKKAASND3FG- 384
Db 213 AAKAVALVLELKGKLDGYALRVPVITGSAITLTFNTKSEVTVESINAAIKAAVGEFGE 272
QY 385 ---YTEDPIVSSDIVGVSYGSLFDATQTKMEVDGSLQVKVSWYDNEMSYTAQLVTLTE 441
Db 273 TLAYSEPLVSTDIVHDSHSGSIIDAGLTKV----SGNTVKVSWYDNEMGYTCQLRLTE 328
QY 442 YFA 444
Db 329 LVA 331

RESULT 15
G3P1_BACSU
ID G3P1_BACSU STANDARD; PRG; 334 AA.
AC P09124;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH) (NAD-
DE dependent glyceraldehyde-3-phosphate dehydrogenase).
GN GAP OR GAP OR BSU33940.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=169 / BD170;
RX MEDLINE=89160255; PubMed=2493629;
RA Viaene A., Dhaese P.;
RT "Sequence of the glyceraldehyde-3-phosphate dehydrogenase gene from
RT Bacillus subtilis."
RL Nucleic Acids Res. 17:1251-1251(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer J., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Chou S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gham S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Gutsepp G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lepidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mael C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche E., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,

```

Search completed: September 15, 2004, 10:24:57
Job time : 27 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2004, 10:18:54 ; Search time 119 Seconds
(without alignments)
1187.833 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278

Sequence: 1 MKKVTGIIILLAVILSAC.....EMSYTAQLVRLVYFAKIAK 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1656.5	72.7	336	2 Q83ZF7	Q83zf7 streptococc
2	1564.5	68.7	336	2 Q8ALW2	Q8alw2 streptococc
3	1564.5	68.7	336	16 Q8E3E8	Q8e3e8 streptococc
4	1564.5	68.7	336	16 Q8DXS8	Q8dxs8 streptococc
5	1559.5	68.5	336	2 Q8KVU6	Q8kvu6 streptococc
6	1544.5	67.8	336	2 Q8GCR7	Q8gcr7 streptococc
7	1535	67.4	335	16 Q97NL1	Q97nl1 streptococc
8	1535	67.4	359	16 Q8CWN6	Q8cwn6 streptococc
9	1534.5	67.4	336	2 Q8KHG1	Q8khg1 streptococc
10	1530.5	67.2	336	2 Q8KVU5	Q8kvu5 streptococc
11	1521	66.8	337	16 Q8DVV3	Q8dvv3 streptococc
12	1513.5	66.4	336	2 Q8VVB9	Q8vvb9 streptococc
13	1459.5	64.1	320	2 Q9L5X6	Q9l5x6 streptococc
14	1410	61.9	309	2 Q9AJT7	Q9ajt7 streptococc
15	1409	61.9	309	2 Q9AJT9	Q9ajt9 streptococc
16	1407	61.8	309	2 Q9AJT4	Q9ajt4 streptococc

17	1405	61.7	309	2 Q9AJT5	Q9ajt5 streptococc
18	1404	61.6	308	2 Q9AJT8	Q9ajt8 streptococc
19	1387	60.9	305	2 Q9AJT6	Q9ajt6 streptococc
20	1357.5	59.6	336	16 Q9CDH4	Q9cdh4 lactococcu
21	1282	56.3	333	16 Q833I8	Q833i8 enterococcu
22	1205.5	52.9	336	16 Q8Y4I1	Q8y4i1 listeria mo
23	1204.5	52.9	336	16 Q928H9	Q928h9 listeria in
24	1175	51.6	332	16 Q8XKT9	Q8xkt9 clostridium
25	1164	51.1	335	16 Q8RFN9	Q8rfn9 fusobacteri
26	1145	50.3	334	16 Q9JW78	Q9jw78 neisseria m
27	1138.5	50.0	336	16 Q8CPY5	Q8cpy5 staphylococ
28	1134	49.8	330	2 Q83UX3	Q83ux3 neisseria g
29	1132.5	49.7	336	16 Q9ZSC5	Q9zsc5 staphylococ
30	1132	49.7	334	16 Q9JX95	Q9jx95 neisseria m
31	1131	49.6	330	2 Q84HZ6	Q84hz6 neisseria g
32	1130	49.6	330	2 Q84HZ4	Q84hz4 neisseria g
33	1129	49.6	335	2 Q93M61	Q93m61 staphylococ
34	1127	49.5	330	2 Q84HZ5	Q84hz5 neisseria g
35	1089	47.8	310	2 Q8L2P7	Q8l2p7 staphylococ
36	1082.5	47.5	311	2 Q8L2R0	Q8l2r0 staphylococ
37	1075	47.2	310	2 Q8L2Q2	Q8l2q2 staphylococ
38	1074	47.1	310	2 Q8L2Q6	Q8l2q6 staphylococ
39	1073	47.1	310	2 Q8L2P6	Q8l2p6 staphylococ
40	1071	47.0	310	2 Q8L2P5	Q8l2p5 staphylococ
41	1070	47.0	311	2 Q8VM79	Q8vm79 rhodospiril
42	1069	46.9	310	2 Q8L2Q3	Q8l2q3 staphylococ
43	1068.5	46.9	311	2 Q8L2P8	Q8l2p8 staphylococ
44	1067	46.8	310	2 Q8L2Q7	Q8l2q7 staphylococ
45	1065	46.8	310	2 Q8L2R7	Q8l2r7 staphylococ

ALIGNMENTS

RESULT 1

Q83ZF7 PRELIMINARY; PRT; 336 AA.

AC Q83ZF7; ID Q83ZF7; DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Glyceraldehyde 3-P dehydrogenase.
GN GAPC.
OS Streptococcus dysgalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1334;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Gasal J.F., Bolton A., Song X.-M., Willson P., Potter A.A.;
RT "Use of the surface proteins GapC and Mig of Streptococcus
RT dysgalactiae as protective antigens against mastitis in non-lactating
RT cows.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF375662; AAP31408.1; .
DR GO; GO:0004365; F-glyceraldehyde-3-phosphate dehydrogenase (p...; IEA.
DR GO; GO:0006096; P-glycolysis; IEA.
DR InterPro; IPR006424; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRfam; TIGR01534; GAPDH-I; 1.
DR PROSITE; P500071; GAPDH; 1.
SQ SEQUENCE 336 AA; 35928 MW; 11828218CF037076 CRC64;

Query Match 72.7%; Score 1656.5; DB 2; Length 336;

Best Local Similarity 79.6%; Pred. No. 2.3e-88; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGFRIGRLAFRIQNVGEVTRINDLTDPNLAHLKYDTTQGRFDGTVEV 87

Db 1 MVVKVINGFGFRIGRLAFRIQNVGEVTRINDLTDPNLAHLKYDTTQGRFDGTVEV 60

QY 88 KEGFEVNGNFIVKSAERDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
 |||||
 Db 61 KEGFEVNGNFIVKSAERDPE-----81
 |||||
 QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 |||||-----NIDWATDGVIEVLE 95
 Db 82 -----
 QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTGTETVIGASCCTTN 267
 |||||
 Db 96 ATGFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTGTETVIGASCCTTN 155
 |||||
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSGAA 327
 |||||
 Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSGAA 215
 |||||
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGVLVTLTKNVSVDEINAAKKAASNDSFGYTE 387
 |||||
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGVLVTLTKNVSVDEINAAKKAASNDSFGYTE 275
 |||||
 QY 388 DPTVSSDIVGVSGLFDATQTKMVEVDSGLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
 |||||
 Db 276 DPTVSSDIVGVSGLFDATQTKMVEVDSGLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
 |||||
 QY 448 K 448
 Db 336 K 336

RESULT 2

Q9ALW2 ID Q9ALW2 PRELIMINARY; PRT; 336 AA.
 AC Q9ALW2;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase.
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J48;
 RA Seifert K.N., Bleiweis A.S., McArthur W.P., Brady L.J.;
 RT "The Group B Streptococcal Surface Antigen Delta is a Glyceraldehyde
 3-Phosphate Dehydrogenase";
 RL Submitted (Jan-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 DEHYDROGENASE FAMILY.
 DR EMBL; AF338416; AAK14387.1; -.
 DR HSSP; P00362; 1GD1.
 DR DR
 DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0016491; P:glycolysis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR000173; GAP_dhdrogenase.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh; 1.
 DR PRINTS; PR00078; G3PDHGRNASE.
 DR TIGRFAMs; TIGR01534; GAPDH-1; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 68.7%; Score 1564.5; DB 2; Length 336;
 Best Local Similarity 74.1%; Pred. No. 5e-83;
 Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGLAFRIQNVGVETRIIDLTPNMLAHLKYDTTQGRFDGTVFV 87
 |||||
 Db 1 MVVKVINGFGRIGLAFRIQNVGVETRIIDLTPNMLAHLKYDTTQGRF-----54
 |||||

QY 88 KEGFEVNGNFIVKSAERDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
 |||||
 Db 55 -----54
 |||||
 QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 |||||-----DGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 95
 Db 55 -----
 QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTGTETVIGASCCTTN 267
 |||||
 Db 96 ATGFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTGTETVIGASCCTTN 155
 |||||
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSGAA 327
 |||||
 Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAANIVPNSGAA 215
 |||||
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGVLVTLTKNVSVDEINAAKKAASNDSFGYTE 387
 |||||
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGVLVTLTKNVSVDEINAAKKAASNDSFGYTE 275
 |||||
 QY 388 DPTVSSDIVGVSGLFDATQTKMVEVDSGLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
 |||||
 Db 276 DPTVSSDIVGVSGLFDATQTKVQVTDGNQLVKVSWYDNEMSYTSQLVRLTLEYFAKIA 335
 |||||
 QY 448 K 448
 Db 336 K 336

RESULT 3

Q8E3E8 ID Q8E3E8 PRELIMINARY; PRT; 336 AA.
 AC Q8E3E8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase.
 GN GBS1811.
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=216495;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NEM316 / Serotype III;
 RX MEDLINE=2242508; PubMed=12354221;
 RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
 RA Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
 RA Kunst F.;
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
 RT invasive neonatal disease.";
 RT Mcl. Microbiol. 45:1499-1513(2002).
 DR EMBL; AL766853; CAD47470.1; -.
 DR SagalList; gbs1811; -.
 DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR006424; GAPDH-1.
 DR InterPro; IPR000173; GAP_dhdrogenase.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh; 1.
 DR PRINTS; PR00078; G3PDHGRNASE.
 DR TIGRFAMs; TIGR01534; GAPDH-1; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 KW Complete proteome.
 SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 68.7%; Score 1564.5; DB 16; Length 336;
 Best Local Similarity 74.1%; Pred. No. 5e-83;
 Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGLAFRIQNVGVETRIIDLTPNMLAHLKYDTTQGRFDGTVFV 87
 |||||
 Db 1 MVVKVINGFGRIGLAFRIQNVGVETRIIDLTPNMLAHLKYDTTQGRF-----54
 |||||

Query Match 68.7% Score 1564.5; DB 16; Length 336;
Best Local Similarity 74.1%; Pred. No. 5e-83;
Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRF 87
DB 1 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRF 54
QY 88 KEGFVNGNFIKVSAREDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVS 147
DB 55 ----- 54
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
DB 55 ----- 95
QY 208 ATSPFAKAEAEKHLHANGAKKVVITAPGGNDVKTIVTNTNHDILDTGTETVISGASCTTN 267
DB 96 ATGFASKEAEQHIHENGAKKVVITAPGGNDVKTIVTNTNHDILDTGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDQMLDGPHEGGDLRRARAGAANIVPNSTGAA 327
DB 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDQMLDGPHEGGDLRRARAGAANIVPNSTGAA 215
QY 328 KAIGLIVIPELNGKLDGAAQRPVPTGVTVELVTLTKNVSVDENINAAKKAANDSFGYTE 387
DB 216 KAIGLIVIPELNGKLDGAAQRPVPTGVTVELVTLTKNVSVDENINAAKKAANDSFGYTE 275
QY 388 DPVSSDIVGYSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAOLVRLTFYFAKIA 447
DB 276 DPVSSDIVGYSGLFDATQTKVQTVDNQNLVKVSWYDNEMSYTSQLVRLTFYFAKIA 335
QY 448 K 448
DB 336 K 336

RESULT 5
Q8KVU6 PRELIMINARY; PRT; 336 AA.
AC Q8KVU6; AC Q8KVU6; PRELIMINARY; PRT; 336 AA.
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
OS Streptococcus uberis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1349;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 9927;
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
RA Potter A.A.;
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT antigens confers protection against heterologous challenge with
RT Streptococcus uberis";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL: AF421900; AAM73771.1; -;
DR GO: GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO: GO:0016491; P:oxidoreductase activity; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR InterPro: IPR006424; GAPDH-I.
DR InterPro: IPR000173; GAP_dhhydrogenase.
DR Pfam: PF00044; gpdh.1.
DR Pfam: PF02800; gpdh.C.1.
DR PRINTS: PR00078; G3PDHGRGNASE.
DR TIGRFAMS: TIGR01534; GAPDH-I; 1.
DR PROSITE: PS00071; GAPDH; 1.
KW Oxidoreductase.

DB 1 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRF 54
QY 88 KEGFVNGNFIKVSAREDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVS 147
DB 55 ----- 54
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
DB 55 ----- 95
QY 208 ATSPFAKAEAEKHLHANGAKKVVITAPGGNDVKTIVTNTNHDILDTGTETVISGASCTTN 267
DB 96 ATGFASKEAEQHIHENGAKKVVITAPGGNDVKTIVTNTNHDILDTGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDQMLDGPHEGGDLRRARAGAANIVPNSTGAA 327
DB 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDQMLDGPHEGGDLRRARAGAANIVPNSTGAA 215
QY 328 KAIGLIVIPELNGKLDGAAQRPVPTGVTVELVTLTKNVSVDENINAAKKAANDSFGYTE 387
DB 216 KAIGLIVIPELNGKLDGAAQRPVPTGVTVELVTLTKNVSVDENINAAKKAANDSFGYTE 275
QY 388 DPVSSDIVGYSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAOLVRLTFYFAKIA 447
DB 276 DPVSSDIVGYSGLFDATQTKVQTVDNQNLVKVSWYDNEMSYTSQLVRLTFYFAKIA 335
QY 448 K 448
DB 336 K 336

RESULT 4
Q8DXS8 PRELIMINARY; PRT; 336 AA.
AC Q8DXS8; AC Q8DXS8; PRELIMINARY; PRT; 336 AA.
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
GN GAP OR SAGI768.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Masiognani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarcelli M., Moxa M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL: AE014272; AAN00631.1; -;
DR TIGR; SAGI768; -;
DR GO: GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR InterPro: IPR006424; GAPDH-I.
DR InterPro: IPR000173; GAP_dhhydrogenase.
DR Pfam: PF00044; gpdh.1.
DR Pfam: PF02800; gpdh.C.1.
DR PRINTS: PR00078; G3PDHGRGNASE.
DR TIGRFAMS: TIGR01534; GAPDH-I; 1.
DR PROSITE: PS00071; GAPDH; 1.
KW Complete proteome.
SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

SQ SEQUENCE 336 AA; 35906 MW; 755D74AA4548E61D4 CRC64;

Query Match 68.5%; Score 1559.5; DB 2; Length 336;
 Best Local Similarity 73.9%; Pred. No. 9, 7e-83;
 Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETVTRINDLTDENMLAHLKLYDTTQGRFDGTVEV 87
 DB 1 MVKVGINGFGRIGRLAFRRIONVEGVETVTRINDLTDENMLAHLKLYDTTQGRFDGTVEV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIVLEALEGTVEVKDGGDFVNGKFIKVSAAE 147
 DB 61 KGGFEVNGNFIKVSARDPE----- 81

QY 148 KDEQIDWATDGVIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIVLE 207
 DB 82 -----NIDWATDGVIVLE 95

QY 208 ATSPFAKAEAEKHLHANGAKKVVITAPGGNDVTVVFNTHDILDGTETVISGASCTTN 267
 DB 96 ATGFFAKAEAEKHLHANGAKKVVITAPGGDDVTVVFNTHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDQMDLDPHRRGGDLRRARAGAAANIVPNSGTAA 327
 DB 156 CLAPMAKALQDNFGVKQGLMTTHAYTGDQMDLDPHRRGGDLRRARAGASNIIVPNSGTAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLVDKNVSVDEINAAKKAASNDSFGYTE 387
 DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLVAVLEKETSVEEINAAKKAASNDSFGYTE 275

QY 388 DPVSSDIVGVSGLSFDATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
 DB 276 DPVSSDIIGWAYGSLFDATQTKVQTDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335

QY 448 K 448
 DB 336 K 336

RESULT 6
 Q8GCR7
 ID Q8GCR7 PRELIMINARY; PRT; 336 AA.
 AC Q8GCR7;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12).
 OS Streptococcus suis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1307;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S735;
 RA Brassard J., Gottschalk M., Quessey S.;
 RT "Cloning and purification of Streptococcus suis serotype 2
 RT glyceralddehyde-3-phosphate dehydrogenase."
 RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF167026; AAN86058.1; -
 DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . ; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR006424; GAPDH-I.
 DR InterPro; IPR000173; GAP_dhdrogenase.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh; 1.
 DR PRINTS; PR00078; G3PDHDRGNASE.
 DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 336 AA; 35825 MW; 1DBB9B1A492DCF59 CRC64;

Query Match 67.8%; Score 1544.5; DB 2; Length 336;

Best Local Similarity 72.9%; Pred. No. 7, 2e-82;
 Matches 307; Conservative 16; Mismatches 13; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETVTRINDLTDENMLAHLKLYDTTQGRFDGTVEV 87
 DB 1 MVKVGINGFGRIGRLAFRRIONVEGVETVTRINDLTDENMLAHLKLYDTTQGRF----- 54

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIVLEALEGTVEVKDGGDFVNGKFIKVSAAE 147
 DB 55 ----- 54

QY 148 KDEQIDWATDGVIVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIVLE 207
 DB 55 -----DGTVEVKDGGFEVNGKFKVSAEREPANIDWATDGVIVLE 95

QY 208 ATSPFAKAEAEKHLHANGAKKVVITAPGGNDVTVVFNTHDILDGTETVISGASCTTN 267
 DB 96 ATGFFAKAEAEKHLHANGAKKVVITAPGGNDVTVVFNTHDILDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDQMDLDPHRRGGDLRRARAGAAANIVPNSGTAA 327
 DB 156 CLAPMAKALHDAFGVQKGLMTTHAYTGDQMDLDPHRRGGDLRRARAGAAANIVPNSGTAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLVDKNVSVDEINAAKKAASNDSFGYTE 387
 DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLVDKKTAAEVNAAKKAATESFGYTE 275

QY 388 DPVSSDIVGVSGLSFDATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
 DB 276 DQVSSDIVGISFGSLFDATQTKVIEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335

QY 448 K 448
 DB 336 K 336

RESULT 7
 Q97NLI
 ID Q97NLI PRELIMINARY; PRT; 335 AA.
 AC Q97NLI;
 DT 01-OCT-2001 (T-EMBLrel. 18, Created)
 DT 01-OCT-2001 (T-EMBLrel. 18, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12).
 GN SP2012 OR GAPDH.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC BAA-334 / TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.B., Lewis M.R., Radune D.,
 RA Holtzapfel E., Knouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae";
 RL Science 293:498-506(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC11733;
 RA Bergmann S., Hammerschmidt S.;
 RT "Identification of pneumococcal GAPDH as plasmin(ogen)-binding
 RT protein";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE007430; AAK76079.1; -
 DR EMBL; AJ505822; CAD44376.1; -.

RA PIR; F95235; F95235.
RA HSP; P00354; 3GPD.
RA TIGR; SP2012; -
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 335 AA; 35856 MW; DA483CEA423E747B CRC64;

Query Match 67.4%; Score 1535; DB 16; Length 335;
Best Local Similarity 73.6%; Pred. No. 2.5e-81;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNLAHLKYDTTQGRFDGTV 87
DB 1 MVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNLAHLKYDTTQGRFDGTV 60
QY 88 KEGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVEVDKGGDFVNGKFIKVS 147
DB 61 KEGFEVNGNFIKVSARDPE----- 81
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGFEVNGQFVKVSAREPANIDWATDGVLE 207
DB 82 -----QIDWATDGVLE 95
QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVFNTNHDLDGTETVISGASCTTN 267
DB 96 ATGFFAKKEAEKHL-KGGAKKVVITAPGNDVKTVFNTNHDLDGTETVISGASCTTN 154
QY 268 CLAPMAKALHDHAFQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAIVPNSTGAA 327
DB 155 CLAPMAKALQDNFVVEGLMTTHAYTGDQMLDGHRRGDLRRARAGAAIVPNSTGAA 214
QY 328 KAIGLVIPELNGKLDGSAORVPVPTGVSVELVTLDKNVSVDENAAKKAASND 387
DB 215 KAIGLVIPELNGKLDGSAORVPVPTGVSVELVTLVAVLEKNVTVDEVNAAMKAASNES 274
QY 388 DPTVSSDIVGVSYSGLFDATQTKWMEVDGSLVKKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
DB 275 DPTVSSDIVGVSYSGLFDATQTKVLDVGKQLVKKVSWYDNEMSYTAQLVRLTLEYFAKIA 334
QY 448 K 448
DB 335 K 335

RESULT 8
ID Q8CWN6 PRELIMINARY; PRT; 359 AA.
AC Q8CWN6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (Phosphorylating)
DE (EC 1.2.1.12).
GN GAPA OR SPRI825.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Knoja H., Kraft A.R., Lagace R.E., P.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,

RA McAhren S.M., McHenney M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris P.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
RA Sun P.-M., Winkler M.B., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL; AE008547; AAL00628.1; -
DR F.R.; G98099; G98099.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 359 AA; 38763 MW; 61EF0E375B330B60 CRC64;

Query Match 67.4%; Score 1535; DB 16; Length 359;
Best Local Similarity 73.6%; Pred. No. 2.8e-81;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNLAHLKYDTTQGRFDGTV 87
DB 25 MVKVGINGFGRIGRLAFRRIONVEGVETRLNDLTPNLAHLKYDTTQGRFDGTV 84
QY 88 KEGFEVNGNFIKVSARDPENIDWATDGVLEALEGTVEVDKGGDFVNGKFIKVS 147
DB 85 KEGFEVNGNFIKVSARDPE----- 105
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGFEVNGQFVKVSAREPANIDWATDGVLE 207
DB 106 -----QIDWATDGVLE 119
QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVFNTNHDLDGTETVISGASCTTN 267
DB 120 ATGFFAKKEAEKHL-KGGAKKVVITAPGNDVKTVFNTNHDLDGTETVISGASCTTN 178
QY 268 CLAPMAKALHDHAFQKGLMTTHAYTGDQMLDGHRRGDLRRARAGAAIVPNSTGAA 327
DB 179 CLAPMAKALQDNFVVEGLMTTHAYTGDQMLDGHRRGDLRRARAGAAIVPNSTGAA 238
QY 328 KAIGLVIPELNGKLDGSAORVPVPTGVSVELVTLDKNVSVDENAAKKAASND 387
DB 239 KAIGLVIPELNGKLDGSAORVPVPTGVSVELVTLVAVLEKNVTVDEVNAAMKAASNES 298
QY 388 DPTVSSDIVGVSYSGLFDATQTKWMEVDGSLVKKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
DB 299 DPTVSSDIVGVSYSGLFDATQTKVLDVGKQLVKKVSWYDNEMSYTAQLVRLTLEYFAKIA 358
QY 448 K 448
DB 359 K 359

RESULT 9
ID Q8KHG1 PRELIMINARY; PRT; 336 AA.
AC Q8KHG1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
DE Streptococcus agalactiae, and
OS Streptococcus iniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311, 1346;
RN [1]

```

RP SEQUENCE FROM N.A.
RC SPECIES=S.agalactiae, and S.iniae; STRAIN=ATCC 27541, and 9117;
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
RA Potter A.A.;
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT antigens confers protection against heterologous challenge with
RT Streptococcus uberis.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF421899; AAM73770.1; -.
DR EMBL; AF421902; AAM73773.1; -.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .); IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 35723 MW; ABAB9E14P3ED1111 CRC64;

Query Match 67.4%; Score 1534.5; DB 2; Length 336;
Best Local Similarity 72.7%; Pred. No. 2.7e-81;
Matches 306; Conservative 17; Mismatches 13; Indels 85; Gaps 1;

QY 28 MVVKGINGFGRIGRLAFRIQNVGVEVTRINDLTPNNLAHLKYDITQGRPDGTVEV 87
Db 1 MVVKGINGFGRIGRLAFRIQNVGVEVTRINDLTPNNLAHLKYDITQGRPDGTVEV 60
QY 88 KEGGFVNGNFIKVSAERDPENIDWATDGVIEVLEALGTVGVKGGDFVNGFKVSAE 147
Db 61 KGGFVNGS----- 70
QY 148 KOPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 71 -----FVKVSAEREPANIDWATDGVIEVLE 95
QY 208 ATSFPAKKEAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDGTETVIGSAGCTTN 267
Db 96 ATGFFAKAAAEQHIHANGAKKVITAPGNDVKTVVFNTHDILDGTETVIGSAGCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSGTAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSGTAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVTLVTLVTLVTLVTLVTLVTLVTLVTLV 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVTLVTLVTLVTLVTLVTLVTLVTLVTLV 275
QY 388 DPVSSDIVGVSGLFDATQTKVMEVDGSQLVKKVSVWYDNMSYTAQLVRLTLEYFAKIA 447
Db 276 DAIVSSDIVGVSGLFDATQTKVQTVDNQNLVKKVSVWYDNMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 10
Q8KVU5 PRELIMINARY; PRT; 336 AA.
AC Q8KVU5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
OS Streptococcus parauberis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

```

```

OX NCBI_TaxID=1348;
RN [1] SEQUENCE FROM N.A.
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
RA Potter A.A.;
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT antigens confers protection against heterologous challenge with
RT Streptococcus uberis.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF421901; AAM73772.1; -.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .); IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh C; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 36037 MW; 311647C2549AC9E CRC64;

Query Match 67.2%; Score 1530.5; DB 2; Length 336;
Best Local Similarity 73.4%; Pred. No. 4.6e-81;
Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2;

QY 28 MVVKGINGFGRIGRLAFRIQNVGVEVTRINDLTPNNLAHLKYDITQGRPDGTVEV 87
Db 1 MVVKGINGFGRIGRLAFRIQNVGVEVTRINDLTPNNLAHLKYDITQGRPDGTVEV 55
QY 88 KEGGFVNGNFIKVSAERDPENIDWATDGVIEVLEALGTVGVKGGDFVNGFKVSAE 147
Db 56 -----GTVGVKGGDFVNGFKVSAE 77
QY 148 KOPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 78 KOPEQ-----IDWATDGVIEVLE 95
QY 208 ATSFPAKKEAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDGTETVIGSAGCTTN 267
Db 96 ATGFFAKAAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDGTETVIGSAGCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSGTAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSGTAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVTLVTLVTLVTLVTLVTLVTLVTLVTLV 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVTLVTLVTLVTLVTLVTLVTLVTLVTLV 275
QY 388 DPVSSDIVGVSGLFDATQTKVMEVDGSQLVKKVSVWYDNMSYTAQLVRLTLEYFAKIA 447
Db 276 DPVSSDIVGVSGLFDATQTKVQTVDNQNLVKKVSVWYDNMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 11
Q8DVV3 PRELIMINARY; PRT; 337 AA.
AC Q8DVV3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Extracellular glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12).
GN GAPC OR SMU.360.
OS Streptococcus mutans.

```

OS	Streptococcus thermophilus.	OS	Streptococcus thermophilus.
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC	Streptococcus.	OC	Streptococcus.
OX	NCBI_TaxID=1309;	OX	NCBI_TaxID=1309;
EN	[1]	EN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	STRAIN=UAI59 / ATCC 700610 / Serotype C;	RC	STRAIN=UAI59 / ATCC 700610 / Serotype C;
RA	MEDLINE=22295063; PubMed=12397186;	RA	MEDLINE=22295063; PubMed=12397186;
RA	Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,	RA	Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA	Carlson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,	RA	Carlson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA	Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.,	RA	Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.,
RT	"Genome sequence of Streptococcus mutans UAI59, a cariogenic dental	RT	"Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT	pathogen.";	RT	pathogen.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).	RL	Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
DR	EMBL; AE014883; AAN58118.1; --	DR	EMBL; AE014883; AAN58118.1; --
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. .; IEA.	DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. .; IEA.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.	DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0006096; P:glycolysis; IEA.	DR	GO; GO:0006096; P:glycolysis; IEA.
DR	InterPro; IPR006424; GAPDH-I.	DR	InterPro; IPR006424; GAPDH-I.
DR	Pfam; PF00044; gpdh; 1.	DR	Pfam; PF00044; gpdh; 1.
DR	PRINTS; PR00078; G3PDHGRNASE.	DR	PRINTS; PR00078; G3PDHGRNASE.
DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.	DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR	PROSITE; PS00071; GAPDH; 1.	DR	PROSITE; PS00071; GAPDH; 1.
KW	Oxidoreductase; Complete proteome.	KW	Oxidoreductase; Complete proteome.
SQ	SEQUENCE 337 AA; 36068 MW; 42BFE20365963C22 CRC64;	SQ	SEQUENCE 337 AA; 36068 MW; 42BFE20365963C22 CRC64;
Query Match 66.8%; Score 1521; DB 16; Length 337;			
Best Local Similarity 71.8%; Pred. No. 1.7e-80;			
Matches 303; Conservative 18; Mismatches 15; Indels 86; Gaps 2;			
QY	28 MVKVGINGFGRIAPRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87	QY	28 MVKVGINGFGRIAPRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db	1 MVKVGINGFGRIAPRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60	Db	1 MVKVGINGFGRIAPRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY	88 KEGGFVNGNFKVSAERDPENIDWATDGEIVLEALEGTVKDGDFWNGKFIKVS 147	QY	88 KEGGFVNGNFKVSAERDPENIDWATDGEIVLEALEGTVKDGDFWNGKFIKVS 147
Db	61 KEGGFVNGNFKVSAERDPE----- 81	Db	61 KEGGFVNGNFKVSAERDPE----- 81
QY	148 KDPEQIDWATDGEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGEIV 207	QY	148 KDPEQIDWATDGEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGEIV 207
Db	82 -----QIDWAADGVEIVLE 95	Db	82 -----QIDWAADGVEIVLE 95
QY	208 ATSFPAKKEAAEKHLHAN -GAKKVVITAPGNDVKTVFNTHDIDGTETVSGASCTT 266	QY	208 ATSFPAKKEAAEKHLHAN -GAKKVVITAPGNDVKTVFNTHDIDGTETVSGASCTT 266
Db	96 ATGFPAKAAAEKHLHANGAKKVVITAPGNDIKTIVFNTHDVLDTETVSGASCTT 155	Db	96 ATGFPAKAAAEKHLHANGAKKVVITAPGNDIKTIVFNTHDVLDTETVSGASCTT 155
QY	267 NCLAPMAKALHDAFGIOKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGA 326	QY	267 NCLAPMAKALHDAFGIOKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGA 326
Db	156 NCLAPMAKALHDFSIKEGLMTTHAYTGDMVLDGPHRGDLRRARAGAAANIVPNSTGA 215	Db	156 NCLAPMAKALHDFSIKEGLMTTHAYTGDMVLDGPHRGDLRRARAGAAANIVPNSTGA 215
QY	327 AKAGLVIPELNGKLDGAAQRPVPTGVTVELVTLTKNVSDVINAAKKAASNDSPGYT 386	QY	327 AKAGLVIPELNGKLDGAAQRPVPTGVTVELVTLTKNVSDVINAAKKAASNDSPGYT 386
Db	216 AKAGLVIPELNGKLDGAAQRPVPTGVTVELVTLTKNVSDVINAAKKAASNDSPGYT 275	Db	216 AKAGLVIPELNGKLDGAAQRPVPTGVTVELVTLTKNVSDVINAAKKAASNDSPGYT 275
QY	387 EDPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTEYPAKI 446	QY	387 EDPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTEYPAKI 446
Db	276 EDPIVSSDIVGMSGLFDATQTKVLDVKGQLVKVSVWYDNEMSYTSQLVRLTEYPAKI 335	Db	276 EDPIVSSDIVGMSGLFDATQTKVLDVKGQLVKVSVWYDNEMSYTSQLVRLTEYPAKI 335
QY	447 AK 448	QY	447 AK 448
Db	336 AK 337	Db	336 AK 337
RESULT 12			
ID	Q8VVB9	ID	Q8VVB9
AC	PRELIMINARY; PRT; 336 AA.	AC	PRELIMINARY; PRT; 336 AA.
DT	01-MAR-2002 (TrEMBLrel. 20, Created)	DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2003 (TrEMBLrel. 20, Last sequence update)	DT	01-MAR-2003 (TrEMBLrel. 20, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Glyceraldehyde-3-phosphate dehydrogenase.	DE	Glyceraldehyde-3-phosphate dehydrogenase.
GN	GAPDH.	GN	GAPDH.

OS	Streptococcus thermophilus.	OS	Streptococcus thermophilus.
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC	Streptococcus.	OC	Streptococcus.
OX	NCBI_TaxID=1308;	OX	NCBI_TaxID=1308;
EN	[1]	EN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	STRAIN=LMG18311;	RC	STRAIN=LMG18311;
RA	van den Bogaard P.T.C., Kleerebezem M., Hols P., Crispie F.,	RA	van den Bogaard P.T.C., Kleerebezem M., Hols P., Crispie F.,
RA	Kuipers O.P., de Vos W.M.;	RA	Kuipers O.P., de Vos W.M.;
RT	"Modulation of Glycolysis by Lactose Availability in Streptococcus	RT	"Modulation of Glycolysis by Lactose Availability in Streptococcus
RT	thermophilus.";	RT	thermophilus.";
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.	RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC	-!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE	CC	-!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC	DEHYDROGENASE FAMILY.	CC	DEHYDROGENASE FAMILY.
DR	EMBL; AF442551; AAL35377.1; --	DR	EMBL; AF442551; AAL35377.1; --
DR	HSSP; P00354; 3GPD.	DR	HSSP; P00354; 3GPD.
DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. .; IEA.	DR	GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. .; IEA.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.	DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0006096; P:glycolysis; IEA.	DR	GO; GO:0006096; P:glycolysis; IEA.
DR	InterPro; IPR006424; GAPDH-I.	DR	InterPro; IPR006424; GAPDH-I.
DR	InterPro; IPR000173; GAP_dhydrogenase.	DR	InterPro; IPR000173; GAP_dhydrogenase.
DR	Pfam; PF00044; gpdh; 1.	DR	Pfam; PF00044; gpdh; 1.
DR	Pfam; PF02800; gpdh_C; 1.	DR	Pfam; PF02800; gpdh_C; 1.
DR	PRINTS; PR00078; G3PDHGRNASE.	DR	PRINTS; PR00078; G3PDHGRNASE.
DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.	DR	TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR	PROSITE; PS00071; GAPDH; 1.	DR	PROSITE; PS00071; GAPDH; 1.
KW	Oxidoreductase.	KW	Oxidoreductase.
SQ	SEQUENCE 336 AA; 36026 MW; 52C1F25F3A7E0230 CRC64;	SQ	SEQUENCE 336 AA; 36026 MW; 52C1F25F3A7E0230 CRC64;
Query Match 66.4%; Score 1513.5; DB 2; Length 336;			
Best Local Similarity 71.7%; Pred. No. 4.5e-80;			
Matches 302; Conservative 18; Mismatches 16; Indels 85; Gaps 1;			
QY	28 MVKVGINGFGRIAPRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87	QY	28 MVKVGINGFGRIAPRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 87
Db	1 MVKVGINGFGRIAPRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60	Db	1 MVKVGINGFGRIAPRIQNVGEVTRINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY	88 KEGGFVNGNFKVSAERDPENIDWATDGEIVLEALEGTVKDGDFWNGKFIKVS 147	QY	88 KEGGFVNGNFKVSAERDPENIDWATDGEIVLEALEGTVKDGDFWNGKFIKVS 147
Db	61 KEGGFVNGNFKVSAERDPE----- 81	Db	61 KEGGFVNGNFKVSAERDPE----- 81
QY	148 KDPEQIDWATDGEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGEIV 207	QY	148 KDPEQIDWATDGEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGEIV 207
Db	82 -----QIDWANDGVEIVLE 95	Db	82 -----QIDWANDGVEIVLE 95
QY	208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTHDIDGTETVSGASCTTN 267	QY	208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTHDIDGTETVSGASCTTN 267
Db	96 ATGFPTKKVLAEEKHLHPGAKKVVITAPGNDVKTVFNTHDIDGTETVSGASCTTN 155	Db	96 ATGFPTKKVLAEEKHLHPGAKKVVITAPGNDVKTVFNTHDIDGTETVSGASCTTN 155
QY	268 CLAPMAKALHDAFGIOKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327	QY	268 CLAPMAKALHDAFGIOKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
Db	156 CLAPMAKALNDNFGIVEGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 215	Db	156 CLAPMAKALNDNFGIVEGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 215
QY	328 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVTLTKNVSDVINAAKKAASNDSPGYTE 387	QY	328 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVTLTKNVSDVINAAKKAASNDSPGYTE 387
Db	216 KAIGLVIPELNGKLDGSAQRPVPTGVTVELVAVCEKNVTVDENAAKKAATNESGYTE 275	Db	216 KAIGLVIPELNGKLDGSAQRPVPTGVTVELVAVCEKNVTVDENAAKKAATNESGYTE 275
QY	388 DPTVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTEYPAKIA 447	QY	388 DPTVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSVWYDNEMSYTAQLVRLTEYPAKIA 447
Db	276 DPTVSSDIVGMSYGLFDATQTKVLDVKGQLVKVSVWYDNEMSYTSQLVRLTEYPAKIA 335	Db	276 DPTVSSDIVGMSYGLFDATQTKVLDVKGQLVKVSVWYDNEMSYTSQLVRLTEYPAKIA 335
QY	448 K 448	QY	448 K 448
Db	336 K 336	Db	336 K 336
RESULT 13			
ID	Q9LSX6	ID	Q9LSX6
AC	PRELIMINARY; PRT; 320 AA.	AC	PRELIMINARY; PRT; 320 AA.
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	DT	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Extracellular glyceraldehyde-3-phosphate dehydrogenase
 DE (Fragment).
 OS Streptococcus gordonii.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1302;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nelson D., Goldstein J.M., Boatright K., Harty D.W.S., Cook S.L.,
 RA Hickman P.J., Potempa J., Travis J., Mayo J.A.;
 RT "Purification and characterization of an extracellular
 RT glyceraldehyde-3-phosphate dehydrogenase from Streptococcus sanguis
 RT and cloning of the gene encoding this enzyme."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDAJ databases.
 CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 CC EMBL; AF247678; AAF64063.1; --
 DR HSSP; P00362; IGD1
 DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p...); IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR006424; GAPDH-I.
 DR InterPro; IPR000173; GAP_dhhydrogenase.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh; 1.
 DR PRINTS; PR00078; G3PDHGRGNASE.
 DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 DR Oxidoreductase.
 FT NON TER 1 1
 FT NON TER 320 320
 SQ SEQUENCE 320 AA; 33968 MW; 6FFCC18B01E91F0 CRC64;

Query Match 64.1%; Score 1459.5; DB 2; Length 320;
 Best Local Similarity 71.9%; Pred. No. 5,7e-77;
 Matches 291; Conservative 14; Mismatches 15; Indels 85; Gaps 1;

QY 31 KVGINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEVKEG 90
 Db 1 KVGINGFGRIGRLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEVKEG 60
 QY 91 GFVNGNFKVSAREPENIDWATDGVIEALEGTVVEKGGDFVNGKFIKVSAREKDP 150
 Db 61 GFVNGKFKVSAREDE ----- 78
 QY 151 EQIDWATDGVIEILEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEILEATS 210
 Db 79 -----NIDWANDGVIEILEATG 95
 QY 211 FFAKKEAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVISGASCTTNCLA 270
 Db 96 FFAKKEAAEKHLHANGAKKVVITAPGSDVKTVVFNTHDILDGTETVISGASCTTNCLA 155
 QY 271 PMAKALHDAFGI:QKGLMTTHAYTGDMILDPHRRGGDI:RRAPAGAAANIVPNSGTGAAKAI 330
 Db 156 PMAKALQDNFVGVGLMTTHAYTGDMILDPHRRGGDI:RRAPAGAAANIVPNSGTGAAKAI 215
 QY 331 GLVIEPGLNGKLGAAQRPVPTGVTGLVTLVDKNVSVDEINAAKAAASNDGSGYGTEDPI 390
 Db 216 GLVIEPGLNGKLGAAQRPVPTGVTGLVTLVDKNVSVDEINAAKAAASNESYGYGTEDPI 275
 QY 391 VSSDIVGVSYGSLFDATQTKVMEVDGSQLVKVSVYDNEMSYTAQ 435
 Db 276 VSSDIVGVSYGSLFDATQTKVLDVKGQLVKVSVYDNEMSYTAQ 320

RESULT 14
 Q9AJT7
 ID Q9AJT7 PRELIMINARY; PRT; 309 AA.
 AC Q9AJT7;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
 DE (Fragment).
 GN GAP.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-clinical isolate;
 RA Amezcaga M.R., Carter P.E., Cash P., McKenzie H.;
 RT "Sequence variation in mef(A) and gap genes in M-phenotype
 RT erythromycin resistant Streptococcus pneumoniae."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DDAJ databases.
 CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 CC EMBL; AJ292048; CAC27448.1; --
 DR HSSP; P00362; IGD1
 DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p...); IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR006424; GAPDH-I.
 DR InterPro; IPR000173; GAP_dhhydrogenase.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh; 1.
 DR PRINTS; PR00078; G3PDHGRGNASE.
 DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 DR Oxidoreductase.
 FT NON TER 1 1
 FT NON TER 309 309
 SQ SEQUENCE 309 AA; 32930 MW; 89E45C3BEA6A0528 CRC64;

Query Match 61.9%; Score 1410; DB 2; Length 309;
 Best Local Similarity 71.9%; Pred. No. 4e-74;
 Matches 284; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 36 GFGRIGLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEVKEGFEVN 95
 Db 1 GFGRIGLAFRRIONVEGVETTRINDLTPNMLAHLKYDTTQGRFDGTVEVKEGFEVN 60
 QY 96 GFNIKVSAREPENIDWATDGVIEALEGTVVEKGGDFVNGKFIKVSAREKDPQIDW 155
 Db 61 GFNIKVSAREDE ----- 73
 QY 156 ATDGVIEILEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEILEATSFFAKK 215
 Db 74 -----QIDWATDGVIEILEATGFFAKK 95
 QY 216 EAAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKA 275
 Db 96 EAAEKHL-KGGAKKVVITAPGNDVKTVVFNTHDILDGTETVISGASCTTNCLAPMAKA 154
 QY 276 LHDAFGIOKGLMTTHAYTGDMILDPHRRGGDI:RRAPAGAAANIVPNSGTGAAKAIGLVIP 335
 Db 155 LQDNFVGVGLMTTHAYTGDMILDPHRRGGDI:RRAPAGAAANIVPNSGTGAAKAIGLVIP 214
 QY 336 ELNGLKLGAAQRPVPTGVTGLVTLVDKNVSVDEINAAKAAASNDGSGYGTEDPIVSSDI 395
 Db 215 ELNGLKLGAAQRPVPTGVTGLVTLVDKNVSVDEINAAKAAASNESYGTEDPIVSSDI 274
 QY 396 VGVSYGSLFDATQTKVMEVDGSQLVKVSVYDNEM 430
 Db 275 VGVSYGSLFDATQTKVLDVKGQLVKVSVYDNEM 309

RESULT 15
 Q9AJT9
 ID Q9AJT9 PRELIMINARY; PRT; 309 AA.
 AC Q9AJT9;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
DE (Fragment).
GN GAP.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=clinical isolate;
RA Amezaga M.R., Carter P.E., Cash P., McKenzie H.;
RT "Sequence variation in mef(A) and gap genes in M-phenotype
RT erythromycin resistant Streptococcus pneumoniae.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AJ292046; CAC27446.1; -.
DR HSSP; P00362; IGD1.
DR GO; GO:0004365; P:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhydrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh_C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 1
FT NON_TER 309
SQ SEQUENCE 309 AA; 32916 MW; 39E9CC3BEA6F8959 CRC64;

Query Match 61.9%; Score 1409; DB 2; Length 309;
Best Local Similarity 71.6%; Pred. No. 4.6e-74;
Matches 283; Conservative 13; Mismatches 13; Indels 86; Gaps 2;

QY 36 GFGRIQLAFRRIQNVGVEVTRINDLTPNMLAHLLKYDTTQGRFDGTVEVKEGFEVN 95
Db 1 GFGRIQLAFRRIQNVGVEVTRINDLTPNMLAHLLKYDTTQGRFDGTVEVKEGFEVN 60

QY 96 GNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAREKDPQIDW 155
Db 61 GKEVKVSAERDPE----- 73

QY 156 ATDGVLEIDGTVEVKEGFEVNGQVKVSAEREPANIDWATDGVLEIVLEATSPFAK 215
Db 74 -----QIDWATDGVLEIVLEATSPFAK 95

QY 216 EAAEKHLHANGAKKVITAPGNDVKTVPNTNHDILDTGTETVISGASCTNCLAPMAKA 275
Db 96 EAAEKHL-KGAKKVITAPGNDVKTVPNTNHDVLDGTETVISGASCTNCLAPMAKA 154

QY 276 LHDAFGIQGLMTTHAYTGDMILDGPHRGDLRRRAGAANTVPNSTGAAGAIGLVIP 335
Db 155 LQDNFVGVEGLMTTHAYTGDMILDGPHRGDLRRRAGAANTVPNSTGAAGAIGLVIP 214

QY 336 ELNGKLDGAARVPVPGSVTELVTLLDKVSVDEINAAKKAASNDSPGYTEDPIVSSDI 395
Db 215 ELNGKLDGAARVPVPGSVTELVTLLDKVSVDEINAAKKAASNDSPGYTEDPIVSSDI 274

QY 396 VGVSYGSLFDATQTKVMEVDGSQLVKVSVYDNEM 430
Db 275 VGVSYGSLFDATQTKVLDVDGKQLVKVSVYDNEM 309

Search completed: September 15, 2004, 10:27:05
Job time : 123 secs

This Page Blank (uspto)